

# 1. PARTICLE DETECTORS AT ACCELERATORS

This is an abridgment of the discussion given in the full *Review of Particle Physics* (the “full *Review*”); the equation and reference numbering corresponds to that version. The quoted numbers are usually based on typical devices, and should be regarded only as rough approximations for new designs. A more detailed discussion of detectors can be found in Refs. 1 and 44.

## 1.1. Introduction

This review summarizes the detector technologies employed at accelerator particle physics experiments. Several of these detectors are also used in a non-accelerator context and examples of such applications will be provided. The detector techniques which are specific to non-accelerator particle physics experiments are the subject of Chap. 32. More detailed discussions of detectors and their underlying physics can be found in books by Ferbel [1], Kleinknecht [2], Knoll [3], Green [4], Leroy & Rancoita [5], and Grupen [6].

In Table 1.1 are given typical resolutions and deadtimes of common charged particle detectors. The quoted numbers are usually based on typical devices, and should be regarded only as rough approximations for new designs. The spatial resolution refers to the intrinsic detector resolution, i.e. without multiple scattering. We note that analog detector readout can provide better spatial resolution than digital readout by measuring the deposited charge in neighboring channels. Quoted ranges attempt to be representative of both possibilities. The time resolution is defined by how accurately the time at which a particle crossed the detector can be determined. The deadtime is the minimum separation in time between two resolved hits on the same channel. Typical performance of calorimetry and particle identification are provided in the relevant sections below.

**Table 1.1:** Typical resolutions and deadtimes of common charged particle detectors. Revised November 2011.

Detector Type	Intrinsic Spatial Resolution (rms)	Time Resolution	Dead Time
Resistive plate chamber	$\lesssim 10$ mm	1–2 ns	—
Streamer chamber	$300\ \mu\text{m}^a$	$2\ \mu\text{s}$	100 ms
Liquid argon drift [7]	$\sim 175\text{--}450\ \mu\text{m}$	$\sim 200$ ns	$\sim 2\ \mu\text{s}$
Scintillation tracker	$\sim 100\ \mu\text{m}$	$100\ \text{ps}/n^b$	10 ns
Bubble chamber	$10\text{--}150\ \mu\text{m}$	1 ms	$50\ \text{ms}^c$
Proportional chamber	$50\text{--}100\ \mu\text{m}^d$	2 ns	20–200 ns
Drift chamber	$50\text{--}100\ \mu\text{m}$	$2\ \text{ns}^a$	20–100 ns
Micro-pattern gas detectors	$30\text{--}40\ \mu\text{m}$	$< 10$ ns	10–100 ns
Silicon strip	pitch/(3 to 7) <sup>a</sup>	few ns <sup>a</sup>	$\lesssim 50\ \text{ns}^a$
Silicon pixel	$\lesssim 10\ \mu\text{m}$	few ns <sup>a</sup>	$\lesssim 50\ \text{ns}^a$
Emulsion	$1\ \mu\text{m}$	—	—

<sup>a</sup> See full *Review* for qualifications and assumptions.

## 2 1. Detectors at accelerators

### 1.2. Photon detectors

Updated August 2011 by D. Chakraborty (Northern Illinois U) and T. Sumiyoshi (Tokyo Metro U).

Most detectors in high-energy, nuclear, and astrophysics rely on the detection of photons in or near the visible range,  $100\text{ nm} \lesssim \lambda \lesssim 1000\text{ nm}$ , or  $E \approx$  a few eV. This range covers scintillation and Cherenkov radiation as well as the light detected in many astronomical observations.

Generally, photodetection involves generating a detectable electrical signal proportional to the (usually very small) number of incident photons. The process involves three distinct steps:

1. Generation of a primary photoelectron or electron-hole ( $e$ - $h$ ) pair by an incident photon by the photoelectric or photoconductive effect,
2. Amplification of the p.e. signal to detectable levels by one or more multiplicative bombardment steps and/or an avalanche process (usually), and,
3. Collection of the secondary electrons to form the electrical signal.

The important characteristics of a photodetector include the following in statistical averages:

1. Quantum efficiency (QE or  $\epsilon_Q$ ): the number of primary photoelectrons generated per incident photon ( $0 \leq \epsilon_Q \leq 1$ ; in silicon more than one  $e$ - $h$  pair per incident photon can be generated for  $\lambda \lesssim 165\text{ nm}$ ),
2. Collection efficiency (CE or  $\epsilon_C$ ): the overall acceptance factor other than the generation of photoelectrons ( $0 \leq \epsilon_C \leq 1$ ),
3. Gain ( $G$ ): the number of electrons collected for each photoelectron generated,
4. Dark current or dark noise: the electrical signal when there is no photon,
5. Energy resolution: electronic noise (ENC or  $N_e$ ) and statistical fluctuations in the amplification process compound the Poisson distribution of  $n_\gamma$  photons from a given source:

$$\frac{\sigma(E)}{\langle E \rangle} = \sqrt{\frac{f_N}{n_\gamma \epsilon_Q \epsilon_C} + \left( \frac{N_e}{G n_\gamma \epsilon_Q \epsilon_C} \right)^2}, \quad (1.1)$$

where  $f_N$ , or the excess noise factor (ENF), is the contribution to the energy distribution variance due to amplification statistics [9],

6. Dynamic range: the maximum signal available from the detector (this is usually expressed in units of the response to noise-equivalent power, or NEP, which is the optical input power that produces a signal-to-noise ratio of 1),
7. Time dependence of the response: this includes the transit time, which is the time between the arrival of the photon and the electrical pulse, and the transit time spread, which contributes to the pulse rise time and width, and
8. Rate capability: inversely proportional to the time needed, after the arrival of one photon, to get ready to receive the next.

The QE is a strong function of the photon wavelength ( $\lambda$ ), and is usually quoted at maximum, together with a range of  $\lambda$  where the QE is comparable to its maximum. Spatial uniformity and linearity with respect to the number of photons are highly desirable in a photodetector's response.

## 1. Detectors at accelerators 3

**Table 1.2:** Representative characteristics of some photodetectors commonly used in particle physics. The time resolution of the devices listed here vary in the 10–2000 ps range.

Type	$\lambda$ (nm)	$\epsilon_Q \epsilon_C$	Gain	Risetime (ns)	Area (mm <sup>2</sup> )	1-p.e noise (Hz)	HV (V)	Price (USD)
PMT*	115–1700	0.15–0.25	$10^3$ – $10^7$	0.7–10	$10^2$ – $10^5$	$10$ – $10^4$	500–3000	100–5000
MCP*	100–650	0.01–0.10	$10^3$ – $10^7$	0.15–0.3	$10^2$ – $10^4$	0.1–200	500–3500	10–6000
HPD*	115–850	0.1–0.3	$10^3$ – $10^4$	7	$10^2$ – $10^5$	$10$ – $10^3$	$\sim 2 \times 10^4$	$\sim 600$
GPM*	115–500	0.15–0.3	$10^3$ – $10^6$	$O(0.1)$	$O(10)$	$10$ – $10^3$	300–2000	$O(10)$
APD	300–1700	$\sim 0.7$	$10$ – $10^8$	$O(1)$	$10$ – $10^3$	$1$ – $10^3$	400–1400	$O(100)$
PPD	320–900	0.15–0.3	$10^5$ – $10^6$	$\sim 1$	1–10	$O(10^6)$	30–60	$O(100)$
VLPC	500–600	$\sim 0.9$	$\sim 5 \times 10^4$	$\sim 10$	1	$O(10^4)$	$\sim 7$	$\sim 1$

\*These devices often come in multi-anode configurations. In such cases, area, noise, and price are to be considered on a “per readout-channel” basis.

Optimization of these factors involves many trade-offs and vary widely between applications. For example, while a large gain is desirable, attempts to increase the gain for a given device also increases the ENF and after-pulsing (“echos” of the main pulse). In solid-state devices, a higher QE often requires a compromise in the timing properties. In other types, coverage of large areas by focusing increases the transit time spread.

Other important considerations also are highly application-specific. These include the photon flux and wavelength range, the total area to be covered and the efficiency required, the volume available to accommodate the detectors, characteristics of the environment such as chemical composition, temperature, magnetic field, ambient background, as well as ambient radiation of different types and, mode of operation (continuous or triggered), bias (high-voltage) requirements, power consumption, calibration needs, aging, cost, and so on. Several technologies employing different phenomena for the three steps described above, and many variants within each, offer a wide range of solutions to choose from. The salient features of the main technologies and the common variants are described below. Some key characteristics are summarized in Table 1.2.

**1.2.1. Vacuum photodetectors:** Vacuum photodetectors can be broadly subdivided into three types: photomultiplier tubes, microchannel plates, and hybrid photodetectors.

**1.2.1.1. Photomultiplier tubes:** A versatile class of photon detectors, vacuum photomultiplier tubes (PMT) has been employed by a vast majority of all particle physics experiments to date [9]. Both “transmission-” and “reflection-type” PMT’s are widely used. In the former, the photocathode material is deposited on the inside of a transparent window through which the photons enter, while in the latter, the photocathode material rests on a separate surface that the incident photons strike. The cathode material has a low work function, chosen for the wavelength band of interest. When a photon hits the cathode and liberates an electron (the photoelectric effect), the latter is accelerated and guided by electric fields to impinge on a secondary-emission electrode, or dynode, which then emits a few ( $\sim 5$ ) secondary electrons. The multiplication process is repeated typically 10 times in series to generate a sufficient number of electrons,

## 4 1. Detectors at accelerators

which are collected at the anode for delivery to the external circuit. The total gain of a PMT depends on the applied high voltage  $V$  as  $G = AV^{kn}$ , where  $k \approx 0.7\text{--}0.8$  (depending on the dynode material),  $n$  is the number of dynodes in the chain, and  $A$  a constant (which also depends on  $n$ ). Typically,  $G$  is in the range of  $10^5\text{--}10^6$ . Pulse risetimes are usually in the few nanosecond range. With *e.g.* two-level discrimination the effective time resolution can be much better.

A large variety of PMT's, including many just recently developed, covers a wide span of wavelength ranges from infrared (IR) to extreme ultraviolet (XUV) [10]. They are categorized by the window materials, photocathode materials, dynode structures, anode configurations, *etc.* Common window materials are borosilicate glass for IR to near-UV, fused quartz and sapphire ( $\text{Al}_2\text{O}_3$ ) for UV, and  $\text{MgF}_2$  or  $\text{LiF}$  for XUV. The choice of photocathode materials include a variety of mostly Cs- and/or Sb-based compounds such as CsI, CsTe, bi-alkali (SbRbCs, SbKCs), multi-alkali ( $\text{SbNa}_2\text{KCs}$ ), GaAs(Cs), GaAsP, *etc.* Sensitive wavelengths and peak quantum efficiencies for these materials are summarized in Table 1.3. Typical dynode structures used in PMT's are circular cage, line focusing, box and grid, venetian blind, and fine mesh. In some cases, limited spatial resolution can be obtained by using a mosaic of multiple anodes. Fast PMT's with very large windows—measuring up to 508 mm across—have been developed in recent years for detection of Cherenkov radiation in neutrino experiments such as Super-Kamiokande and KamLAND among many others. Specially prepared low-radioactivity glass is used to make these PMT's, and they are also able to withstand the high pressure of the surrounding liquid.

PMT's are vulnerable to magnetic fields—sometimes even the geomagnetic field causes large orientation-dependent gain changes. A high-permeability metal shield is often necessary. However, proximity-focused PMT's, *e.g.* the fine-mesh types, can be used even in a high magnetic field ( $\geq 1$  T) if the electron drift direction is parallel to the field. CMS uses custom-made vacuum phototriodes (VPT) mounted on the back face of projective lead tungstate crystals to detect scintillation light in the endcap sections of its electromagnetic calorimeters, which are inside a 3.8 T superconducting solenoid. A VPT employs a single dynode (thus,  $G \approx 10$ ) placed close to the photocathode, and a mesh anode plane between the two, to help it cope with the strong magnetic field, which is not too unfavorably oriented with respect to the photodetector axis in the endcaps (within  $25^\circ$ ), but where the radiation level is too high for Avalanche Photodiodes (APD's) like those used in the barrel section.

**1.2.1.2. Microchannel plates:** A typical Microchannel plate (MCP) photodetector consists of one or more  $\sim 2$  mm thick glass plates with densely packed  $O(10\text{ }\mu\text{m})$ -diameter cylindrical holes, or “channels”, sitting between the transmission-type photocathode and anode planes, separated by  $O(1\text{ mm})$  gaps. Instead of discrete dynodes, the inner surface of each cylindrical tube serves as a continuous dynode for the entire cascade of multiplicative bombardments initiated by a photoelectron. Gain fluctuations can be minimized by operating in a saturation mode, whence each channel is only capable of a binary output, but the sum of all channel outputs remains proportional to the number of photons received so long as the photon flux is low enough to ensure that the probability of a single channel receiving more than one photon during a single time gate is negligible. MCP's are thin, offer good spatial resolution, have excellent time resolution ( $\sim 20$  ps), and can tolerate random magnetic fields up to 0.1 T and axial fields

up to  $\sim 1$  T. However, they suffer from relatively long recovery time per channel and short lifetime. MCP's are widely employed as image-intensifiers, although not so much in HEP or astrophysics.

**1.2.1.3. Hybrid photon detectors:** Hybrid photon detectors (HPD) combine the sensitivity of a vacuum PMT with the excellent spatial and energy resolutions of a Si sensor [11]. A single photoelectron ejected from the photocathode is accelerated through a potential difference of  $\sim 20$  kV before it impinges on the silicon sensor/anode. The gain nearly equals the maximum number of  $e$ - $h$  pairs that could be created from the entire kinetic energy of the accelerated electron:  $G \approx eV/w$ , where  $e$  is the electronic charge,  $V$  is the applied potential difference, and  $w \approx 3.7$  eV is the mean energy required to create an  $e$ - $h$  pair in Si at room temperature. Since the gain is achieved in a single step, one might expect to have the excellent resolution of a simple Poisson statistic with large mean, but in fact it is even better, thanks to the Fano effect discussed in Sec. 1.7.

Low-noise electronics must be used to read out HPD's if one intends to take advantage of the low fluctuations in gain, *e.g.* when counting small numbers of photons. HPD's can have the same  $\epsilon_Q$   $\epsilon_C$  and window geometries as PMT's and can be segmented down to  $\sim 50$   $\mu$ m. However, they require rather high biases and will not function in a magnetic field. The exception is proximity-focused devices ( $\Rightarrow$  no (de)magnification) in an axial field. With time resolutions of  $\sim 10$  ps and superior rate capability, proximity-focused HPD's can be an alternative to MCP's. Current applications of HPD's include the CMS hadronic calorimeter and the RICH detector in LHCb. Large-size HPD's with sophisticated focusing may be suitable for future water Cherenkov experiments.

Hybrid APD's (HAPD's) add an avalanche multiplication step following the electron bombardment to boost the gain by a factor of  $\sim 50$ . This affords a higher gain and/or lower electrical bias, but also degrades the signal definition.

**Table 1.3:** Properties of photocathode and window materials commonly used in vacuum photodetectors [10].

Photocathode material	$\lambda$ (nm)	Window material	Peak $\epsilon_Q$ ( $\lambda$ /nm)
CsI	115–200	MgF <sub>2</sub>	0.11 (140)
CsTe	115–320	MgF <sub>2</sub>	0.14 (240)
Bi-alkali	300–650	Borosilicate	0.27 (390)
	160–650	Synthetic Silica	0.27 (390)
“Ultra Bi-alkali”	300–650	Borosilicate	0.43 (350)
	160–650	Synthetic Silica	0.43 (350)
Multi-alkali	300–850	Borosilicate	0.20 (360)
	160–850	Synthetic Silica	0.20 (360)
GaAs(Cs)*	160–930	Synthetic Silica	0.23 (280)
GaAsP(Cs)	300–750	Borosilicate	0.50 (500)
InP/InGaAsP <sup>†</sup>	350–1700	Borosilicate	0.01 (1100)

\*Reflection type photocathode is used. <sup>†</sup>Requires cooling to  $\sim -80^\circ\text{C}$ .

## 6 1. Detectors at accelerators

**1.2.2. Gaseous photon detectors:** In gaseous photomultipliers (GPM) a photoelectron in a suitable gas mixture initiates an avalanche in a high-field region, producing a large number of secondary impact-ionization electrons. In principle the charge multiplication and collection processes are identical to those employed in gaseous tracking detectors such as multiwire proportional chambers, micromesh gaseous detectors (Micromegas), or gas electron multipliers (GEM). These are discussed in Sec. 1.6.4.

The devices can be divided into two types depending on the photocathode material. One type uses solid photocathode materials much in the same way as PMT's. Since it is resistant to gas mixtures typically used in tracking chambers, CsI is a common choice. In the other type, photoionization occurs on suitable molecules vaporized and mixed in the drift volume. Most gases have photoionization work functions in excess of 10 eV, which would limit their sensitivity to wavelengths far too short. However, vapors of TMAE (tetrakis dimethyl-amine ethylene) or TEA (tri-ethyl-amine), which have smaller work functions (5.3 eV for TMAE and 7.5 eV for TEA), are suited for XUV photon detection [12]. Since devices like GEM's offer sub-mm spatial resolution, GPM's are often used as position-sensitive photon detectors. They can be made into flat panels to cover large areas ( $O(1 \text{ m}^2)$ ), can operate in high magnetic fields, and are relatively inexpensive. Many of the ring imaging Cherenkov (RICH) detectors to date have used GPM's for the detection of Cherenkov light [13]. Special care must be taken to suppress the photon-feedback process in GPM's. It is also important to maintain high purity of the gas as minute traces of  $\text{O}_2$  can significantly degrade the detection efficiency.

**1.2.3. Solid-state photon detectors:** In a phase of rapid development, solid-state photodetectors are competing with vacuum- or gas-based devices for many existing applications and making way for a multitude of new ones. Compared to traditional vacuum- and gaseous photodetectors, solid-state devices are more compact, lightweight, rugged, tolerant to magnetic fields, and often cheaper. They also allow fine pixelization, are easy to integrate into large systems, and can operate at low electric potentials, while matching or exceeding most performance criteria. They are particularly well suited for detection of  $\gamma$ - and X-rays. Except for applications where coverage of very large areas or dynamic range is required, solid-state detectors are proving to be the better choice. Some hybrid devices attempt to combine the best features of different technologies while applications of nanotechnology are opening up exciting new possibilities.

Silicon photodiodes (PD) are widely used in high-energy physics as particle detectors and in a great number of applications (including solar cells!) as light detectors. The structure is discussed in some detail in Sec. 1.7. In its simplest form, the PD is a reverse-biased  $p$ - $n$  junction. Photons with energies above the indirect bandgap energy (wavelengths shorter than about 1050 nm, depending on the temperature) can create  $e$ - $h$  pairs (the photoconductive effect), which are collected on the  $p$  and  $n$  sides, respectively. Often, as in the PD's used for crystal scintillator readout in CLEO, L3, Belle, BaBar, and GLAST, intrinsic silicon is doped to create a  $p$ - $i$ - $n$  structure. The reverse bias increases the thickness of the depleted region; in the case of these particular detectors, to full depletion at a depth of about 100  $\mu\text{m}$ . Increasing the depletion depth decreases the capacitance (and hence electronic noise) and extends the red response. Quantum efficiency can exceed 90%, but falls toward the red because of the increasing absorption length of light in silicon. The absorption length

reaches  $100\ \mu\text{m}$  at  $985\ \text{nm}$ . However, since  $G = 1$ , amplification is necessary. Optimal low-noise amplifiers are slow, but, even so, noise limits the minimum detectable signal in room-temperature devices to several hundred photons.

Very large arrays containing  $O(10^7)$  of  $O(10\ \mu\text{m}^2)$ -sized photodiodes pixelizing a plane are widely used to photograph all sorts of things from everyday subjects at visible wavelengths to crystal structures with X-rays and astronomical objects from infrared to UV. To limit the number of readout channels, these are made into charge-coupled devices (CCD), where pixel-to-pixel signal transfer takes place over thousands of synchronous cycles with sequential output through shift registers [14]. Thus, high spatial resolution is achieved at the expense of speed and timing precision. Custom-made CCD's have virtually replaced photographic plates and other imagers for astronomy and in spacecraft. Typical QE's exceed 90% over much of the visible spectrum, and "thick" CCD's have useful QE up to  $\lambda = 1\ \mu\text{m}$ . Active Pixel Sensor (APS) arrays with a preamplifier on each pixel and CMOS processing afford higher speeds, but are challenged at longer wavelengths. Much R&D is underway to overcome the limitations of both CCD and CMOS imagers.

In APD's, an exponential cascade of impact ionizations initiated by the original photogenerated  $e$ - $h$  pair under a large reverse-bias voltage leads to an avalanche breakdown [15]. As a result, detectable electrical response can be obtained from low-intensity optical signals down to single photons. Excellent junction uniformity is critical, and a guard ring is generally used as a protection against edge breakdown. Well-designed APD's, such as those used in CMS' crystal-based electromagnetic calorimeter, have achieved  $\epsilon_Q \epsilon_C \approx 0.7$  with sub-ns response time. The sensitive wavelength window and gain depend on the semiconductor used. The gain is typically 10–200 in linear and up to  $10^8$  in Geiger mode of operation. Stability and close monitoring of the operating temperature are important for linear-mode operation, and substantial cooling is often necessary. Position-sensitive APD's use time information at multiple anodes to calculate the hit position.

One of the most promising recent developments in the field is that of devices consisting of large arrays ( $O(10^3)$ ) of tiny APD's packed over a small area ( $O(1\ \text{mm}^2)$ ) and operated in a limited Geiger mode [16]. Among different names used for this class of photodetectors, "PPD" (for "Pixelized Photon Detector") is most widely accepted (formerly "SiPM"). Although each cell only offers a binary output, linearity with respect to the number of photons is achieved by summing the cell outputs in the same way as with a MCP in saturation mode (see above). PPD's are being adopted as the preferred solution for various purposes including medical imaging, *e.g.* positron emission tomography (PET). These compact, rugged, and economical devices allow auto-calibration through decent separation of photoelectron peaks and offer gains of  $O(10^6)$  at a moderate bias voltage ( $\sim 50\ \text{V}$ ). However, the single-photoelectron noise of a PPD, being the logical "or" of  $O(10^3)$  Geiger APD's, is rather large:  $O(1\ \text{MHz}/\text{mm}^2)$  at room temperature. PPD's are particularly well-suited for applications where triggered pulses of several photons are expected over a small area, *e.g.* fiber-guided scintillation light. Intense R&D is expected to lower the noise level and improve radiation hardness, resulting in coverage of larger areas and wider applications. Attempts are being made to combine the fabrication of the sensors and the front-end electronics (ASIC) in the same process with the goal of making PPD's and other finely pixelized solid-state photodetectors extremely easy to use.

## 8 1. Detectors at accelerators

Of late, much R&D has been directed to *p-i-n* diode arrays based on thin polycrystalline diamond films formed by chemical vapor deposition (CVD) on a hot substrate ( $\sim 1000$  K) from a hydrocarbon-containing gas mixture under low pressure ( $\sim 100$  mbar). These devices have maximum sensitivity in the extreme- to moderate-UV region [17]. Many desirable characteristics, including high tolerance to radiation and temperature fluctuations, low dark noise, blindness to most of the solar radiation spectrum, and relatively low cost make them ideal for space-based UV/XUV astronomy, measurement of synchrotron radiation, and luminosity monitoring at (future) lepton collider(s).

Visible-light photon counters (VLPC) utilize the formation of an impurity band only 50 meV below the conduction band in As-doped Si to generate strong ( $G \approx 5 \times 10^4$ ) yet sharp response to single photons with  $\epsilon_Q \approx 0.9$  [18]. The smallness of the band gap considerably reduces the gain dispersion. Only a very small bias ( $\sim 7$  V) is needed, but high sensitivity to infrared photons requires cooling below 10 K. The dark noise increases sharply and exponentially with both temperature and bias. The Run 2 DØ detector used 86000 VLPC's to read the optical signal from its scintillating-fiber tracker and scintillator-strip preshower detectors.

### 1.3. Organic scintillators

Revised August 2011 by Kurtis F. Johnson (FSU).

Organic scintillators are broadly classed into three types, crystalline, liquid, and plastic, all of which utilize the ionization produced by charged particles to generate optical photons, usually in the blue to green wavelength regions [19]. Plastic scintillators are by far the most widely used, liquid organic scintillator is finding increased use, and crystal organic scintillators are practically unused in high-energy physics. Plastic scintillator densities range from 1.03 to 1.20 g cm<sup>-3</sup>. Typical photon yields are about 1 photon per 100 eV of energy deposit [20]. A one-cm-thick scintillator traversed by a minimum-ionizing particle will therefore yield  $\approx 2 \times 10^4$  photons. The resulting photoelectron signal will depend on the collection and transport efficiency of the optical package and the quantum efficiency of the photodetector.

Organic scintillator does not respond linearly to the ionization density. A widely used semi-empirical model by Birks posits that recombination and quenching effects between the excited molecules reduce the light yield [21]. These effects are more pronounced the greater the density of the excited molecules. Birks' formula is

$$\frac{d\mathcal{L}}{dx} = \mathcal{L}_0 \frac{dE/dx}{1 + k_B dE/dx} , \quad (1.2)$$

where  $\mathcal{L}$  is the luminescence,  $\mathcal{L}_0$  is the luminescence at low specific ionization density, and  $k_B$  is Birks' constant, which must be determined for each scintillator by measurement. Decay times are in the ns range; rise times are much faster. The fraction of light emitted during the decay "tail" can depend on the exciting particle. This allows pulse shape discrimination for particle identification. Ease of fabrication into desired shapes and low cost has made plastic scintillator a common detector element. In the form of scintillating fiber it has found widespread use in tracking and calorimetry [23].

Demand for large volume detectors has lead to increased use of liquid organic scintillator, which has the same scintillation mechanism



as plastic scintillator, due to its cost advantage. The containment vessel defines the detector shape; photodetectors or waveshifters may be immersed in the liquid.

### 1.3.2. *Scintillating and wavelength-shifting fibers:*

The clad optical fiber comprising scintillator and wavelength shifter (WLS) is particularly useful [31]. Since the initial demonstration of the scintillating fiber (SCIFI) calorimeter [32], SCIFI techniques have become mainstream [33]. SCIFI calorimeters are fast, dense, radiation hard, and can have leadglass-like resolution. SCIFI trackers can handle high rates and are radiation tolerant, but the low photon yield at the end of a long fiber (see below) forces the use of sensitive photodetectors. WLS scintillator readout of a calorimeter allows a very high level of hermeticity since the solid angle blocked by the fiber on its way to the photodetector is very small. The sensitive region of scintillating fibers can be controlled by splicing them onto clear (non-scintillating/non-WLS) fibers.

A typical configuration would be fibers with a core of polystyrene-based scintillator or WLS (index of refraction  $n = 1.59$ ), surrounded by a cladding of PMMA ( $n = 1.49$ ) a few microns thick, or, for added light capture, with another cladding of fluorinated PMMA with  $n = 1.42$ , for an overall diameter of 0.5 to 1 mm. The fiber is drawn from a boule and great care is taken during production to ensure that the intersurface between the core and the cladding has the highest possible uniformity and quality, so that the signal transmission via total internal reflection has a low loss. The fraction of generated light which is transported down the optical pipe is denoted the capture fraction and is about 6% for the single-clad fiber and 10% for the double-clad fiber. The number of photons from the fiber available at the photodetector is always smaller than desired, and increasing the light yield has proven difficult. A minimum-ionizing particle traversing a high-quality 1 mm diameter fiber perpendicular to its axis will produce fewer than 2000 photons, of which about 200 are captured. Attenuation may eliminate 95% of these photons in a large collider tracker.

A scintillating or WLS fiber is often characterized by its attenuation length, over which the signal is attenuated to  $1/e$  of its original value. Many factors determine the attenuation length, including the importance of re-absorption of emitted photons by the polymer base or dissolved fluors, the level of crystallinity of the base polymer, and the quality of the total internal reflection boundary [34]. Attenuation lengths of several meters are obtained by high quality fibers. However, it should be understood that the attenuation length is not the sole measure of fiber quality. Among other things, it is not constant with distance from the excitation source and it is wavelength dependent.

## 1.4. Inorganic scintillators:

Revised September 2009 by R.-Y. Zhu (California Institute of Technology) and C.L. Woody (BNL).

Inorganic crystals form a class of scintillating materials with much higher densities than organic plastic scintillators (typically  $\sim 4\text{--}8\text{ g/cm}^3$ ) with a variety of different properties for use as scintillation detectors. Due to their high density and high effective atomic number, they can be used in applications where high stopping power or a high conversion efficiency for electrons or photons is required. These include total absorption electromagnetic calorimeters (see Sec. 1.9.1), which consist of a totally active absorber (as opposed to a sampling

## 10 1. Detectors at accelerators

calorimeter), as well as serving as gamma ray detectors over a wide range of energies. Many of these crystals also have very high light output, and can therefore provide excellent energy resolution down to very low energies ( $\sim$  few hundred keV).

Some crystals are intrinsic scintillators in which the luminescence is produced by a part of the crystal lattice itself. However, other crystals require the addition of a dopant, typically fluorescent ions such as thallium (Tl) or cerium (Ce) which is responsible for producing the scintillation light. However, in both cases, the scintillation mechanism is the same. Energy is deposited in the crystal by ionization, either directly by charged particles, or by the conversion of photons into electrons or positrons which subsequently produce ionization. This energy is transferred to the luminescent centers which then radiate scintillation photons. The efficiency  $\eta$  for the conversion of energy deposit in the crystal to scintillation light can be expressed by the relation [35]

$$\eta = \beta \cdot S \cdot Q . \quad (1.3)$$

where  $\beta$  is the efficiency of the energy conversion process,  $S$  is the efficiency of energy transfer to the luminescent center, and  $Q$  is the quantum efficiency of the luminescent center. The value of  $\eta$  ranges between 0.1 and  $\sim 1$  depending on the crystal, and is the main factor in determining the intrinsic light output of the scintillator. In addition, the scintillation decay time is primarily determined by the energy transfer and emission process. The decay time of the scintillator is mainly dominated by the decay time of the luminescent center. For example, in the case of thallium doped sodium iodide (NaI(Tl)), the value of  $\eta$  is  $\sim 0.5$ , which results in a light output  $\sim 40,000$  photons per MeV of energy deposit. This high light output is largely due to the high quantum efficiency of the thallium ion ( $Q \sim 1$ ), but the decay time is rather slow ( $\tau \sim 250$  ns).

Table 28.4 lists the basic properties of some commonly used inorganic crystal scintillators. NaI(Tl) is one of the most common and widely used scintillators, with an emission that is well matched to a bialkali photomultiplier tube, but it is highly hygroscopic and difficult to work with, and has a rather low density. CsI(Tl) has high light yield, an emission that is well matched to solid state photodiodes, and is mechanically robust (high plasticity and resistance to cracking). However, it needs careful surface treatment and is slightly hygroscopic. Compared with CsI(Tl), pure CsI has identical mechanical properties, but faster emission at shorter wavelengths and light output approximately an order of magnitude lower. BaF<sub>2</sub> has a fast component with a sub-nanosecond decay time, and is the fastest known scintillator. However, it also has a slow component with a much longer decay time ( $\sim 630$  ns). Bismuth germanate (Bi<sub>4</sub>Ge<sub>3</sub>O<sub>12</sub> or BGO) has a high density, and consequently a short radiation length  $X_0$  and Molière radius  $R_M$ . BGO's emission is well-matched to the spectral sensitivity of photodiodes, and it is easy to handle and not hygroscopic. Lead tungstate (PbWO<sub>4</sub> or PWO) has a very high density, with a very short  $X_0$  and  $R_M$ , but its intrinsic light yield is rather low.

Cerium doped lutetium oxyorthosilicate (Lu<sub>2</sub>SiO<sub>5</sub>:Ce, or LSO:Ce) [36] and cerium doped lutetium-yttrium oxyorthosilicate (Lu<sub>2(1-x)</sub>Y<sub>2x</sub>SiO<sub>5</sub>, LYSO:Ce) [37] are dense crystal scintillators which have a high light yield and a fast decay time. Only properties of LSO:Ce is listed in Table 28.4 since the properties of LYSO:Ce are similar to that of LSO:Ce except a little lower density than LSO:Ce depending on the yttrium fraction in LYSO:Ce. This material is also

featured with excellent radiation hardness [38], so is expected to be used where extraordinary radiation hardness is required.

Table 28.4 also includes cerium doped lanthanum tri-halides, such as  $\text{LaBr}_3$  [39], which is brighter and faster than  $\text{LSO:Ce}$ , but it is highly hygroscopic and has a lower density. The FWHM energy resolution measured for this material coupled to a PMT with bi-alkali photocathode for 0.662 MeV  $\gamma$ -rays from a  $^{137}\text{Cs}$  source is about 3%, which is the best among all inorganic crystal scintillators. For this reason,  $\text{LaBr}_3$  is expected to be widely used in applications where a good energy resolution for low energy photons are required, such as homeland security.

Beside the crystals listed in Table 28.4, a number of new crystals are being developed that may have potential applications in high energy or nuclear physics. Of particular interest is the family of yttrium and lutetium perovskites, which include YAP ( $\text{YAlO}_3\text{:Ce}$ ) and LuAP ( $\text{LuAlO}_3\text{:Ce}$ ) and their mixed compositions. These have been shown to be linear over a large energy range [40], and have the potential for providing extremely good intrinsic energy resolution. In addition, other fluoride crystals such as  $\text{CeF}_3$  have been shown to provide excellent energy resolution in calorimeter applications.

Aiming at the best jet-mass resolution inorganic scintillators are being investigated for HEP calorimeters with dual readout for both Cherenkov and scintillation light to be used at future linear colliders. These materials may be used for an electromagnetic calorimeter [41] or a homogeneous hadronic calorimetry (HHCAL) detector concept, including both electromagnetic and hadronic parts [42]. Because of the unprecedented volume (70 to 100  $\text{m}^3$ ) foreseen for the HHCAL detector concept the materials must be (1) dense (to minimize the leakage) and (2) cost-effective. It should also be UV transparent (for effective collection of the Cherenkov light) and allow for a clear discrimination between the Cherenkov and scintillation light. The preferred scintillation light is thus at a longer wavelength, and not necessarily bright or fast. Dense crystals, scintillating glasses and ceramics offer a very attractive implementation for this detector concept. Inorganic crystals being investigated are lead fluoride ( $\text{PbF}_2$ ), lead chloride fluoride ( $\text{PbFCl}$ ) and BSO [43].

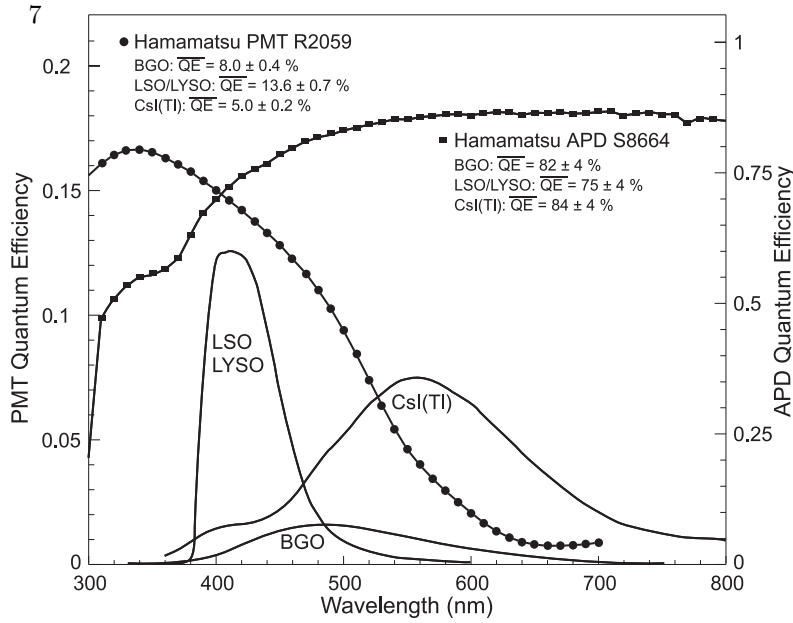
Table 28.4 gives the light output of other crystals relative to  $\text{NaI(Tl)}$  and their dependence to the temperature variations measured for crystal samples of  $1.5 X_0$  cube with a Tyvek paper wrapping and a full end face coupled to a photodetector [44]. The quantum efficiencies of the photodetector is taken out to facilitate a direct comparison of crystal's light output. However, the useful signal produced by a scintillator is usually quoted in terms of the number of photoelectrons per MeV produced by a given photodetector. The relationship between the number of photons/MeV produced and photoelectrons/MeV detected involves the factors for the light collection efficiency  $L$  and the quantum efficiency  $QE$  of the photodetector:

$$N_{\text{p.e.}}/\text{MeV} = L \cdot QE \cdot N_{\gamma}/\text{MeV} \quad (1.4)$$

$L$  includes the transmission of scintillation light within the crystal (*i.e.*, the bulk attenuation length of the material), reflections and scattering from the surfaces, and the size and shape of the crystal. These factors can vary considerably depending on the sample, but can be in the range of  $\sim 10$ –60%. The internal light transmission depends on the intrinsic properties of the material, *e.g.* the density and type of the scattering centers and defects that can produce internal absorption

## 12 1. Detectors at accelerators

within the crystal, and can be highly affected by factors such as radiation damage, as discussed below.



**Figure 1.2:** The quantum efficiencies of two photodetectors, a Hamamatsu R2059 PMT with bi-alkali cathode and a Hamamatsu S8664 avalanche photodiode (APD), are shown as a function of wavelength. Also shown in the figure are emission spectra of three crystal scintillators, BGO, LSO and CsI(Tl), and the numerical values of the emission weighted quantum efficiencies. The area under each emission spectrum is proportional to crystal's light yield.

The quantum efficiency depends on the type of photodetector used to detect the scintillation light, which is typically  $\sim 15\text{--}20\%$  for photomultiplier tubes and  $\sim 70\%$  for silicon photodiodes for visible wavelengths. The quantum efficiency of the detector is usually highly wavelength dependent and should be matched to the particular crystal of interest to give the highest quantum yield at the wavelength corresponding to the peak of the scintillation emission. Fig. 1.2 shows the quantum efficiencies of two photodetectors, a Hamamatsu R2059 PMT with bi-alkali cathode and quartz window and a Hamamatsu S8664 avalanche photodiode (APD) as a function of wavelength. Also shown in the figure are emission spectra of three crystal scintillators, BGO, LSO:Ce/LYSO:Ce and CsI(Tl), and the numerical values of the emission weighted quantum efficiency. The area under each emission spectrum is proportional to crystal's light yield, as shown in Table 28.4, where the quantum efficiencies of the photodetector has been taken out. Results with different photodetectors can be significantly different. For example, the response of CsI(Tl) relative to NaI(Tl) with a standard photomultiplier tube with a bi-alkali photo-cathode, e.g. Hamamatsu R2059, would be 45 rather than 165 because of the photomultiplier's low quantum efficiency at longer wavelengths. For scintillators which emit in the UV, a detector with a quartz window should be used.

For very low energy applications (typically below 1 MeV), non-proportionality of the scintillation light yield may be important. It has been known for a long time that the conversion factor between the energy deposited in a crystal scintillator and the number of photons produced is not constant. It is also known that the energy resolution measured by all crystal scintillators for low energy  $\gamma$ -rays is significantly worse than the contribution from photo-electron statistics alone, indicating an intrinsic contribution from the scintillator itself. Precision measurement using low energy electron beam shows that this non-proportionality is crystal dependent [45]. Recent study on this issue also shows that this effect is also sample dependent even for the same crystal [46]. Further work is therefore needed to fully understand this subject.

One important issue related to the application of a crystal scintillator is its radiation hardness. Stability of its light output, or the ability to track and monitor the variation of its light output in a radiation environment, is required for high resolution and precision calibration [47]. All known crystal scintillators suffer from radiation damage. A common damage phenomenon is the appearance of radiation induced absorption caused by the formation of color centers originated from the impurities or point defects in the crystal. This radiation induced absorption reduces the light attenuation length in the crystal, and hence its light output. For crystals with high defect density, a severe reduction of light attenuation length may cause a distortion of the light response uniformity, leading to a degradation of the energy resolution. Additional radiation damage effects may include a reduced intrinsic scintillation light yield (damage to the luminescent centers) and an increased phosphorescence (afterglow). For crystals to be used in the construction a high precision calorimeter in a radiation environment, its scintillation mechanism must not be damaged and its light attenuation length in the expected radiation environment must be long enough so that its light response uniformity, and thus its energy resolution, does not change [48].

Most of the crystals listed in Table 1.4 have been used in high energy or nuclear physics experiments when the ultimate energy resolution for electrons and photons is desired. Examples are the Crystal Ball NaI(Tl) calorimeter at SPEAR, the L3 BGO calorimeter at LEP, the CLEO CsI(Tl) calorimeter at CESR, the KTeV CsI calorimeter at the Tevatron, the BaBar, BELLE and BES II CsI(Tl) calorimeters at PEP-II, KEK and BEPC III. Because of its high density and relative low cost, PWO calorimeters are widely used by CMS and ALICE at LHC, by CLAS and PrimEx at CEBAF, and by PANDA at GSI. Recently, investigations have been made aiming at using LSO:Ce or LYSO:Ce crystals for future high energy or nuclear physics experiments [38].

## 14 1. Detectors at accelerators

**Table 1.4:** Properties of several inorganic crystal scintillators. Most of the notation is defined in Sec. 6 of this *Review*.

Parameter:	$\rho$	MP	$X_0^*$	$R_M^*$	$dE^*/dx$	$\lambda_I^*$	$\tau_{\text{decay}}$	$\lambda_{\text{max}}$	$n^\ddagger$	Relative output <sup>†</sup>	Hygroscopic?	$d(\text{LY})/dT$
Units:	g/cm <sup>3</sup>	°C	cm	cm	MeV/cm	cm	ns	nm				%/°C <sup>‡</sup>
NaI(Tl)	3.67	651	2.59	4.13	4.8	42.9	245	410	1.85	100	yes	−0.2
BGO	7.13	1050	1.12	2.23	9.0	22.8	300	480	2.15	21	no	−0.9
BaF <sub>2</sub>	4.89	1280	2.03	3.10	6.5	30.7	650 <sup>s</sup>	300 <sup>s</sup>	1.50	36 <sup>s</sup>	no	−1.9 <sup>s</sup>
							0.9 <sup>f</sup>	220 <sup>f</sup>		4.1 <sup>f</sup>		0.1 <sup>f</sup>
CsI(Tl)	4.51	621	1.86	3.57	5.6	39.3	1220	550	1.79	165	slight	0.4
CsI(pure)	4.51	621	1.86	3.57	5.6	39.3	30 <sup>s</sup>	420 <sup>s</sup>	1.95	3.6 <sup>s</sup>	slight	−1.4
							6 <sup>f</sup>	310 <sup>f</sup>		1.1 <sup>f</sup>		
PbWO <sub>4</sub>	8.3	1123	0.89	2.00	10.1	20.7	30 <sup>s</sup>	425 <sup>s</sup>	2.20	0.3 <sup>s</sup>	no	−2.5
							10 <sup>f</sup>	420 <sup>f</sup>		0.077 <sup>f</sup>		
LSO(Ce)	7.40	2050	1.14	2.07	9.6	20.9	40	402	1.82	85	no	−0.2
LaBr <sub>3</sub> (Ce)	5.29	788	1.88	2.85	6.9	30.4	20	356	1.9	130	yes	0.2

\* Numerical values calculated using formulae in this review.

<sup>‡</sup> Refractive index at the wavelength of the emission maximum.

<sup>†</sup> Relative light output measured for samples of 1.5 X<sub>0</sub> cube with a Tyvek paper wrapping and a full end face coupled to a photodetector. The quantum efficiencies of the photodetector are taken out.

<sup>‡</sup> Variation of light yield with temperature evaluated at the room temperature.

<sup>f</sup> Fast component, <sup>s</sup> slow component.  
Revised September 2009 by B.N. Ratcliff (SLAC).

Although devices using Cherenkov radiation are often thought of as only particle identification (PID) detectors, in practice they are used over a broader range of applications including; (1) fast particle counters; (2) hadronic PID; and (3) tracking detectors performing complete event reconstruction. Examples of applications from each category include; (1) the BaBar luminosity detector [49]; (2) the hadronic PID detectors at the B factory detectors—DIRC in BaBar [50] and the aerogel threshold Cherenkov in Belle [51]; and (3) large water Cherenkov counters such as Super-Kamiokande [53]. Cherenkov counters contain two main elements; (1) a radiator through which the charged particle passes, and (2) a photodetector. As Cherenkov radiation is a weak source of photons, light collection and detection must be as efficient as possible. The refractive index  $n$  and the particle's path length through the radiator  $L$  appear in the Cherenkov relations allowing the tuning of these quantities for particular applications.

Cherenkov detectors utilize one or more of the properties of Cherenkov radiation discussed in the Passages of Particles through Matter section (Sec. 30 of this *Review*): the prompt emission of a light pulse; the existence of a velocity threshold for radiation; and the dependence of the Cherenkov cone half-angle  $\theta_c$  and the number of emitted photons on the velocity of the particle and the refractive index of the medium.

The number of photoelectrons ( $N_{\text{p.e.}}$ ) detected in a given device is

$$N_{\text{p.e.}} = L \frac{\alpha^2 z^2}{r_e m_e c^2} \int \epsilon(E) \sin^2 \theta_c(E) dE, \quad (1.5)$$

where  $\epsilon(E)$  is the efficiency for collecting the Cherenkov light and transducing it into photoelectrons, and  $\alpha^2/(r_e m_e c^2) = 370 \text{ cm}^{-1} \text{ eV}^{-1}$ .

The quantities  $\epsilon$  and  $\theta_c$  are functions of the photon energy  $E$ . As the typical energy dependent variation of the index of refraction is modest, a quantity called the *Cherenkov detector quality factor*  $N_0$  can

be defined as

$$N_0 = \frac{\alpha^2 z^2}{r_e m_e c^2} \int \epsilon dE , \quad (1.6)$$

so that, taking  $z = 1$  (the usual case in high-energy physics),

$$N_{\text{p.e.}} \approx L N_0 \langle \sin^2 \theta_c \rangle . \quad (1.7)$$

This definition of the quality factor  $N_0$  is not universal, nor, indeed, very useful for those common situations where  $\epsilon$  factorizes as  $\epsilon = \epsilon_{\text{coll}} \epsilon_{\text{det}}$  with the geometrical photon collection efficiency ( $\epsilon_{\text{coll}}$ ) varying substantially for different tracks while the photon detector efficiency ( $\epsilon_{\text{det}}$ ) remains nearly track independent. In this case, it can be useful to explicitly remove ( $\epsilon_{\text{coll}}$ ) from the definition of  $N_0$ . A typical value of  $N_0$  for a photomultiplier (PMT) detection system working in the visible and near UV, and collecting most of the Cherenkov light, is about  $100 \text{ cm}^{-1}$ . Practical counters, utilizing a variety of different photodetectors, have values ranging between about 30 and  $180 \text{ cm}^{-1}$ . Radiators can be chosen from a variety of transparent materials (Sec. 30 of this *Review* and Table 6.1). In addition to refractive index, the choice requires consideration of factors such as material density, radiation length and radiation hardness, transmission bandwidth, absorption length, chromatic dispersion, optical workability (for solids), availability, and cost. When the momenta of particles to be identified is high, the refractive index must be set close to one, so that the photon yield per unit length is low and a long particle path in the radiator is required. Recently, the gap in refractive index that has traditionally existed between gases and liquid or solid materials has been partially closed with transparent *silica aerogels* with indices that range between about 1.007 and 1.13.

Cherenkov counters may be classified as either *imaging* or *threshold* types, depending on whether they do or do not make use of Cherenkov angle ( $\theta_c$ ) information. Imaging counters may be used to track particles as well as identify them. The recent development of very fast photodetectors such as micro-channel plate PMTs (MCP PMT) (see Sec. 1.2 of this *Review*) also potentially allows very fast Cherenkov based time of flight (TOF) detectors of either class [57].

*Threshold* Cherenkov detectors [54], in their simplest form, make a yes/no decision based on whether the particle is above or below the Cherenkov threshold velocity  $\beta_t = 1/n$ . A straightforward enhancement of such detectors uses the number of observed photoelectrons (or a calibrated pulse height) to discriminate between species or to set probabilities for each particle species [55]. This strategy can increase the momentum range of particle separation by a modest amount (to a momentum some 20% above the threshold momentum of the heavier particle in a typical case).

Careful designs give  $\langle \epsilon_{\text{coll}} \rangle \gtrsim 90\%$ . For a photomultiplier with a typical bialkali cathode,  $\int \epsilon_{\text{det}} dE \approx 0.27 \text{ eV}$ , so that

$$N_{\text{p.e.}}/L \approx 90 \text{ cm}^{-1} \langle \sin^2 \theta_c \rangle \quad (\text{i.e., } N_0 = 90 \text{ cm}^{-1}) . \quad (1.8)$$

Suppose, for example, that  $n$  is chosen so that the threshold for species  $a$  is  $p_t$ ; that is, at this momentum species  $a$  has velocity  $\beta_a = 1/n$ . A second, lighter, species  $b$  with the same momentum has velocity  $\beta_b$ , so  $\cos \theta_c = \beta_a/\beta_b$ , and

$$N_{\text{p.e.}}/L \approx 90 \text{ cm}^{-1} \frac{m_a^2 - m_b^2}{p_t^2 + m_a^2} . \quad (1.9)$$

For  $K/\pi$  separation at  $p = p_t = 1(5) \text{ GeV}/c$ ,  $N_{\text{p.e.}}/L \approx 16(0.8) \text{ cm}^{-1}$  for  $\pi$ 's and (by design) 0 for  $K$ 's.

## 16 1. Detectors at accelerators

For limited path lengths  $N_{\text{p.e.}}$  will usually be small. The overall efficiency of the device is controlled by Poisson fluctuations, which can be especially critical for separation of species where one particle type is dominant. Moreover, the effective number of photoelectrons is often less than the average number calculated above due to additional equivalent noise from the photodetector (see the discussion of the excess noise factor in Sec. 1.2 of this *Review*). It is common to design for at least 10 photoelectrons for the high velocity particle in order to obtain a robust counter. As rejection of the particle that is below threshold depends on *not* seeing a signal, electronic and other background noise can be important. Physics sources of light production for the below threshold particle, such as decay to an above threshold particle or the production of delta rays in the radiator, often limit the separation attainable, and need to be carefully considered. Well designed, modern multi-channel counters, such as the ACC at Belle [51], can attain adequate particle separation performance over a substantial momentum range for essentially the full solid angle of the spectrometer.

*Imaging* counters make the most powerful use of the information available by measuring the ring-correlated angles of emission of the individual Cherenkov photons. Since low-energy photon detectors can measure only the position (and, perhaps, a precise detection time) of the individual Cherenkov photons (not the angles directly), the photons must be “imaged” onto a detector so that their angles can be derived [56]. Typically the optics map the Cherenkov cone onto (a portion of) a distorted “circle” at the photodetector. Though the imaging process is directly analogous to familiar imaging techniques used in telescopes and other optical instruments, there is a somewhat bewildering variety of methods used in a wide variety of counter types with different names. Some of the imaging methods used include (1) focusing by a lens; (2) proximity focusing (i.e., focusing by limiting the emission region of the radiation); and (3) focusing through an aperture (a pinhole). In addition, the prompt Cherenkov emission coupled with the speed of modern photon detectors allows the use of (4) time imaging, a method which is little used in conventional imaging technology. Finally, (5) correlated tracking (and event reconstruction) can be performed in large water counters by combining the individual space position and time of each photon together with the constraint that Cherenkov photons are emitted from each track at the same polar angle (Sec. 32.3.1 of this *Review*).

In a simple model of an imaging PID counter, the fractional error on the particle velocity ( $\delta_\beta$ ) is given by

$$\delta_\beta = \frac{\sigma_\beta}{\beta} = \tan \theta_c \sigma(\theta_c) \quad , \quad (1.10)$$

where

$$\sigma(\theta_c) = \frac{\langle \sigma(\theta_i) \rangle}{\sqrt{N_{\text{p.e.}}}} \oplus C \quad , \quad (1.11)$$

and  $\langle \sigma(\theta_i) \rangle$  is the average single photoelectron resolution, as defined by the optics, detector resolution and the intrinsic chromaticity spread of the radiator index of refraction averaged over the photon detection bandwidth.  $C$  combines a number of other contributions to resolution including, (1) correlated terms such as tracking, alignment, and multiple scattering, (2) hit ambiguities, (3) background hits from random sources, and (4) hits coming from other tracks. The actual separation performance is also limited by physics effects such as decays in flight and particle interactions in the material of the detector. In many practical cases, the performance is limited by these effects.



For a  $\beta \approx 1$  particle of momentum ( $p$ ) well above threshold entering a radiator with index of refraction ( $n$ ), the number of  $\sigma$  separation ( $N_\sigma$ ) between particles of mass  $m_1$  and  $m_2$  is approximately

$$N_\sigma \approx \frac{|m_1^2 - m_2^2|}{2p^2\sigma(\theta_c)\sqrt{n^2 - 1}}. \quad (1.12)$$

In practical counters, the angular resolution term  $\sigma(\theta_c)$  varies between about 0.1 and 5 mrad depending on the size, radiator, and photodetector type of the particular counter. The range of momenta over which a particular counter can separate particle species extends from the point at which the number of photons emitted becomes sufficient for the counter to operate efficiently as a threshold device ( $\sim 20\%$  above the threshold for the lighter species) to the value in the imaging region given by the equation above. For example, for  $\sigma(\theta_c) = 2\text{mrad}$ , a fused silica radiator ( $n = 1.474$ ), or a fluorocarbon gas radiator ( $\text{C}_5\text{F}_{12}$ ,  $n = 1.0017$ ), would separate  $\pi/K$ 's from the threshold region starting around 0.15(3) GeV/ $c$  through the imaging region up to about 4.2(18) GeV/ $c$  at better than  $3\sigma$ .

Many different imaging counters have been built during the last several decades [57]. Among the earliest examples of this class of counters are the very limited acceptance Differential Cherenkov detectors, designed for particle selection in high momentum beam lines. These devices use optical focusing and/or geometrical masking to select particles having velocities in a specified region. With careful design, a velocity resolution of  $\sigma_\beta/\beta \approx 10^{-4}$ – $10^{-5}$  can be obtained [54].

Practical multi-track Ring-Imaging Cherenkov detectors (generically called RICH counters) are a more recent development. RICH counters are sometimes further classified by ‘generations’ that differ based on historical timing, performance, design, and photodetection techniques.

Prototypical examples of first generation RICH counters are those used in the DELPHI and SLD detectors at the LEP and SLC Z factory  $e^+e^-$  colliders [57]. They have both liquid ( $\text{C}_6\text{F}_{14}$ ,  $n = 1.276$ ) and gas ( $\text{C}_5\text{F}_{12}$ ,  $n = 1.0017$ ) radiators, the former being proximity imaged with the latter using mirrors. The phototransducers are a TPC/wire-chamber combination. They are made sensitive to photons by doping the TPC gas (usually, ethane/methane) with  $\sim 0.05\%$  TMAE (tetrakis(dimethylamino)ethylene). Great attention to detail is required, (1) to avoid absorbing the UV photons to which TMAE is sensitive, (2) to avoid absorbing the single photoelectrons as they drift in the long TPC, and (3) to keep the chemically active TMAE vapor from interacting with materials in the system. In spite of their unforgiving operational characteristics, these counters attained good  $e/\pi/K/p$  separation over wide momentum ranges (from about 0.25 to 20 GeV/ $c$ ) during several years of operation at LEP and SLC. Related but smaller acceptance devices include the OMEGA RICH at the CERN SPS, and the RICH in the balloon-borne CAPRICE detector [57].

Later generation counters [57] generally operate at much higher rates, with more detection channels, than the first generation detectors just described. They also utilize faster, more forgiving photon detectors, covering different photon detection bandwidths. Radiator choices have broadened to include materials such as lithium fluoride, fused silica, and aerogel. Vacuum based photodetection systems (*e.g.*, single or multi anode PMTs, MCP PMTs, or hybrid photodiodes (HPD)) have become increasingly common (see Sec. 1.2 of this *Review*). They handle high rates, and can be used with a wide choice

## 18 1. Detectors at accelerators

of radiators. Examples include (1) the SELEX RICH at Fermilab, which mirror focuses the Cherenkov photons from a neon radiator onto a camera array made of  $\sim 2000$  PMTs to separate hadrons over a wide momentum range (to well above  $200 \text{ GeV}/c$  for heavy hadrons); (2) the HERMES RICH at HERA, which mirror focuses photons from  $\text{C}_4\text{F}_{10}$  ( $n = 1.00137$ ) and aerogel ( $n = 1.0304$ ) radiators within the same volume onto a PMT camera array to separate hadrons in the momentum range from 2 to  $15 \text{ GeV}/c$ ; and (3) the LHCb detector now being brought into operation at the LHC. It uses two separate counters. One volume, like HERMES, contains two radiators (aerogel and  $\text{C}_4\text{F}_{10}$ ) while the second volume contains  $\text{CF}_4$ . Photons are mirror focused onto detector arrays of HPDs to cover a  $\pi/K$  separation momentum range between 1 and  $150 \text{ GeV}/c$ .

Other fast detection systems that use solid cesium iodide (CsI) photocathodes or triethylamine (TEA) doping in proportional chambers are useful with certain radiator types and geometries. Examples include (1) the CLEO-III RICH at CESR that uses a LiF radiator with TEA doped proportional chambers; (2) the ALICE detector at the LHC that uses proximity focused liquid ( $\text{C}_6\text{F}_{14}$  radiators and solid CSI photocathodes (similar photodetectors have been used for several years by the HADES and COMPASS detectors), and the hadron blind detector (HBD) in the PHENIX detector at RHIC that couples a low index  $\text{CF}_4$  radiator to a photodetector based on electron multiplier (GEM) chambers with reflective CSI photocathodes [57].

A DIRC (Detection [of] Internally Reflected Cherenkov [light]) is a distinctive, compact RICH subtype first used in the BaBar detector [52]. A DIRC “inverts” the usual RICH principle for use of light from the radiator by collecting and imaging the total internally reflected light rather than the transmitted light. It utilizes the optical material of the radiator in two ways, simultaneously; first as a Cherenkov radiator, and second, as a light pipe. The magnitudes of the photon angles are preserved during transport by the flat, rectangular cross section radiators, allowing the photons to be efficiently transported to a detector outside the path of the particle where they may be imaged in up to three independent dimensions (the usual two in space and, due to the long photon paths lengths, one in time). Because the index of refraction in the radiator is large ( $\sim 1.48$  for fused silica), the momentum range with good  $\pi/K$  separation is rather low. The BaBar DIRC range extends up to  $\sim 4 \text{ GeV}/c$ . It is plausible, but difficult, to extend it up to about  $10 \text{ GeV}/c$  with an improved design. New DIRC detectors are being developed that take advantage of the new, very fast, pixelated photodetectors becoming available, such as flat panel PMTs and MCP PMTs. They typically utilize either time imaging or mirror focused optics, or both, leading not only to a precision measurement of the Cherenkov angle, but in some cases, to a precise measurement of the particle time of flight, and/or to correction of the chromatic dispersion in the radiator. Examples include (1) the time of propagation (TOP) counter being developed for the BELLE-II upgrade at KEKB which emphasizes precision timing for both Cherenkov imaging and TOF; (2) the full 3-dimensional imaging FDIRC for the SuperB detector at the Italian SuperB collider which uses precision timing not only for improving the angle reconstruction and TOF, but also to correct the chromatic dispersion; and (3) the DIRCs being developed for the PANDA detector at FAIR that use elegant focusing optics and fast timing [57].

## 1.6. Gaseous detectors

**1.6.1. Energy loss and charge transport in gases:** Revised March 2010 by F. Sauli (CERN) and M. Titov (CEA Saclay).

Gas-filled detectors localize the ionization produced by charged particles, generally after charge multiplication. The statistics of ionization processes having asymmetries in the ionization trails, affect the coordinate determination deduced from the measurement of drift time, or of the center of gravity of the collected charge. For thin gas layers, the width of the energy loss distribution can be larger than its average, requiring multiple sample or truncated mean analysis to achieve good particle identification. In the truncated mean method for calculating  $\langle dE/dx \rangle$ , the ionization measurements along the track length are broken into many samples and then a fixed fraction of high-side (and sometimes also low-side) values are rejected [58].

The energy loss of charged particles and photons in matter is discussed in Sec. 30. Table 1.5 provides values of relevant parameters in some commonly used gases at NTP (normal temperature, 20° C, and pressure, 1 atm) for unit-charge minimum-ionizing particles (MIPs) [59–65]. Values often differ, depending on the source, so those in the table should be taken only as approximate. For different conditions and for mixtures, and neglecting internal energy transfer processes (*e.g.*, Penning effect), one can scale the density,  $N_P$ , and  $N_T$  with temperature and pressure assuming a perfect gas law.

**Table 1.5:** Properties of noble and molecular gases at normal temperature and pressure (NTP: 20° C, one atm).  $E_X$ ,  $E_I$ : first excitation, ionization energy;  $W_I$ : average energy per ion pair;  $dE/dx|_{\min}$ ,  $N_P$ ,  $N_T$ : differential energy loss, primary and total number of electron-ion pairs per cm, for unit charge minimum ionizing particles.

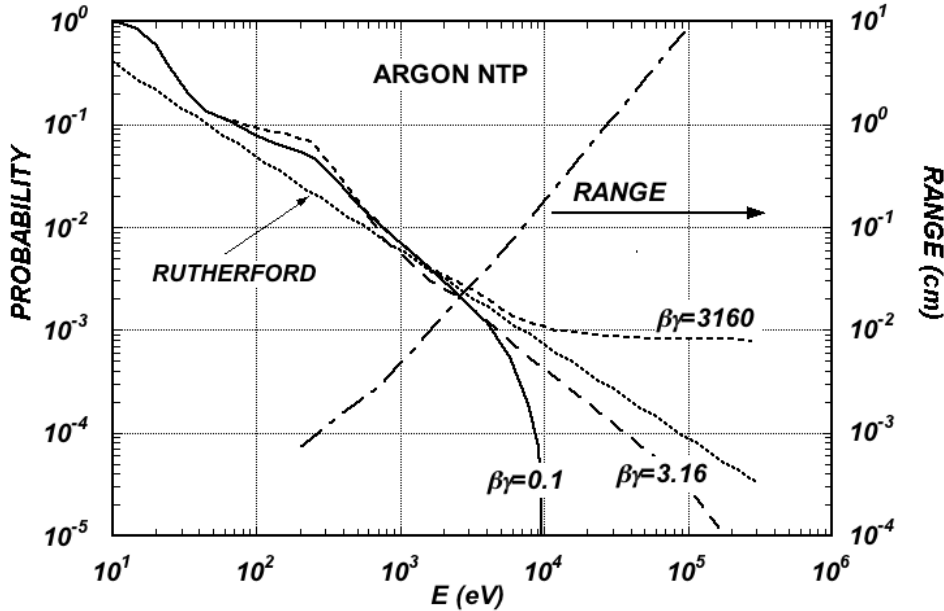
Gas	Density, mg cm <sup>-3</sup>	$E_x$ eV	$E_I$ eV	$W_I$ eV	$dE/dx _{\min}$ keV cm <sup>-1</sup>	$N_P$ cm <sup>-1</sup>	$N_T$ cm <sup>-1</sup>
He	0.179	19.8	24.6	41.3	0.32	3.5	8
Ne	0.839	16.7	21.6	37	1.45	13	40
Ar	1.66	11.6	15.7	26	2.53	25	97
Xe	5.495	8.4	12.1	22	6.87	41	312
CH <sub>4</sub>	0.667	8.8	12.6	30	1.61	28	54
C <sub>2</sub> H <sub>6</sub>	1.26	8.2	11.5	26	2.91	48	112
iC <sub>4</sub> H <sub>10</sub>	2.49	6.5	10.6	26	5.67	90	220
CO <sub>2</sub>	1.84	7.0	13.8	34	3.35	35	100
CF <sub>4</sub>	3.78	10.0	16.0	54	6.38	63	120

When an ionizing particle passes through the gas it creates electron-ion pairs, but often the ejected electrons have sufficient energy to further ionize the medium. As shown in Table 1.5, the total number of electron-ion pairs ( $N_T$ ) is usually a few times larger than the number of primaries ( $N_P$ ).

The probability for a released electron to have an energy  $E$  or larger follows an approximate  $1/E^2$  dependence (Rutherford law), shown in Fig. 1.3 for Ar/CH<sub>4</sub> at NTP (dotted line, left scale). More detailed estimates taking into account the electronic structure of the medium are shown in the figure, for three values of the particle velocity

## 20 1. Detectors at accelerators

factor  $\beta\gamma$  [60]. The dot-dashed line provides, on the right scale, the practical range of electrons (including scattering) of energy  $E$ . As an example, about 0.6% of released electrons have 1 keV or more energy, substantially increasing the ionization loss rate. The practical range of 1 keV electrons in argon (dot-dashed line, right scale) is  $70\ \mu\text{m}$  and this can contribute to the error in the coordinate determination.



**Figure 1.3:** Probability of single collisions in which released electrons have an energy  $E$  or larger (left scale) and practical range of electrons in Ar/CH<sub>4</sub> (P10) at NTP (dot-dashed curve, right scale) [60].

The number of electron-ion pairs per primary ionization, or cluster size, has an exponentially decreasing probability; for argon, there is about 1% probability for primary clusters to contain ten or more electron-ion pairs [61].

Once released in the gas, and under the influence of an applied electric field, electrons and ions drift in opposite directions and diffuse towards the electrodes. The scattering cross section is determined by the details of atomic and molecular structure. Therefore, the drift velocity and diffusion of electrons depend very strongly on the nature of the gas, specifically on the inelastic cross-section involving the rotational and vibrational levels of molecules. In noble gases, the inelastic cross section is zero below excitation and ionization thresholds. Large drift velocities are achieved by adding polyatomic gases (usually CH<sub>4</sub>, CO<sub>2</sub>, or CF<sub>4</sub>) having large inelastic cross sections at moderate energies, which results in “cooling” electrons into the energy range of the Ramsauer-Townsend minimum (at  $\sim 0.5$  eV) of the elastic cross-section of argon. The reduction in both the total electron scattering cross-section and the electron energy results in a large increase of electron drift velocity (for a compilation of electron-molecule cross sections see Ref. 62). Another principal role of the polyatomic gas is to absorb the ultraviolet photons emitted by the excited noble gas atoms. Extensive collections of experimental

data [63] and theoretical calculations based on transport theory [64] permit estimates of drift and diffusion properties in pure gases and their mixtures. In a simple approximation, gas kinetic theory provides the drift velocity  $v$  as a function of the mean collision time  $\tau$  and the electric field  $E$ :  $v = eE\tau/m_e$  (Townsend's expression). Values of drift velocity and diffusion for some commonly used gases at NTP are given in Fig. 1.4 and Fig. 1.5. These have been computed with the MAGBOLTZ program [65]. For different conditions, the horizontal axis must be scaled inversely with the gas density. Standard deviations for longitudinal ( $\sigma_L$ ) and transverse diffusion ( $\sigma_T$ ) are given for one cm of drift, and scale with the square root of the drift distance. Since the collection time is inversely proportional to the drift velocity, diffusion is less in gases such as  $\text{CF}_4$  that have high drift velocities. In the presence of an external magnetic field, the Lorentz force acting on electrons between collisions deflects the drifting electrons and modifies the drift properties. The electron trajectories, velocities and diffusion parameters can be computed with MAGBOLTZ. A simple theory, the friction force model, provides an expression for the vector drift velocity  $\mathbf{v}$  as a function of electric and magnetic field vectors  $\mathbf{E}$  and  $\mathbf{B}$ , of the Larmor frequency  $\omega = eB/m_e$ , and of the mean collision time  $\tau$ :

$$\mathbf{v} = \frac{e}{m_e} \frac{\tau}{1 + \omega^2 \tau^2} \left( \mathbf{E} + \frac{\omega \tau}{B} (\mathbf{E} \times \mathbf{B}) + \frac{\omega^2 \tau^2}{B^2} (\mathbf{E} \cdot \mathbf{B}) \mathbf{B} \right) \quad (1.13)$$

To a good approximation, and for moderate fields, one can assume that the energy of the electrons is not affected by  $B$ , and use for  $\tau$  the values deduced from the drift velocity at  $B = 0$  (the Townsend expression). For  $\mathbf{E}$  perpendicular to  $\mathbf{B}$ , the drift angle to the relative to the electric field vector is  $\tan \theta_B = \omega \tau$  and  $v = (E/B)(\omega \tau / \sqrt{1 + \omega^2 \tau^2})$ . For parallel electric and magnetic fields, drift velocity and longitudinal diffusion are not affected, while the transverse diffusion can be strongly reduced:  $\sigma_T(B) = \sigma_T(B = 0) / \sqrt{1 + \omega^2 \tau^2}$ . The dotted line in Fig. 1.5 represents  $\sigma_T$  for the classic Ar/ $\text{CH}_4$  (90:10) mixture at 4 T. Large values of  $\omega \tau \sim 20$  at 5 T are consistent with the measurement of diffusion coefficient in Ar/ $\text{CF}_4$ /iC<sub>4</sub>H<sub>10</sub> (95:3:2). This reduction is exploited in time projection chambers (Sec. 1.6.5) to improve spatial resolution.

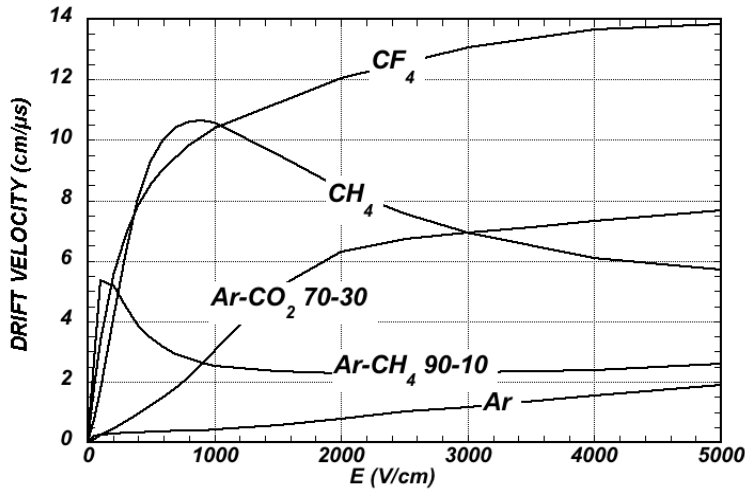


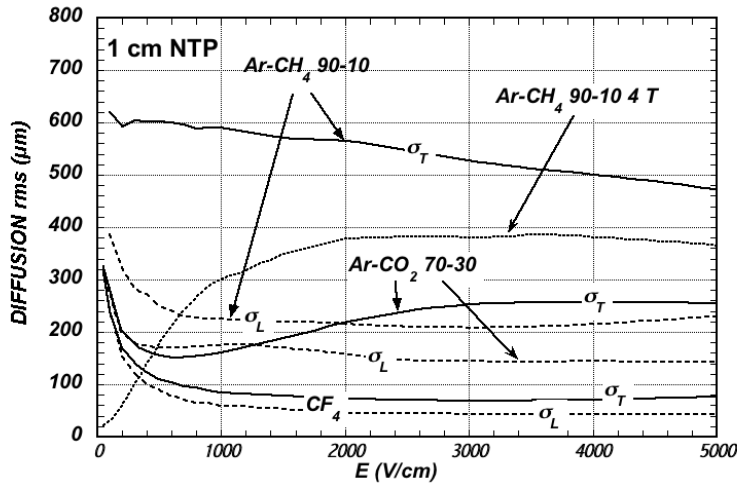
Figure 1.4: Computed electron drift velocity as a function of electric field in several gases at NTP and  $B = 0$  [65].

## 22 1. Detectors at accelerators

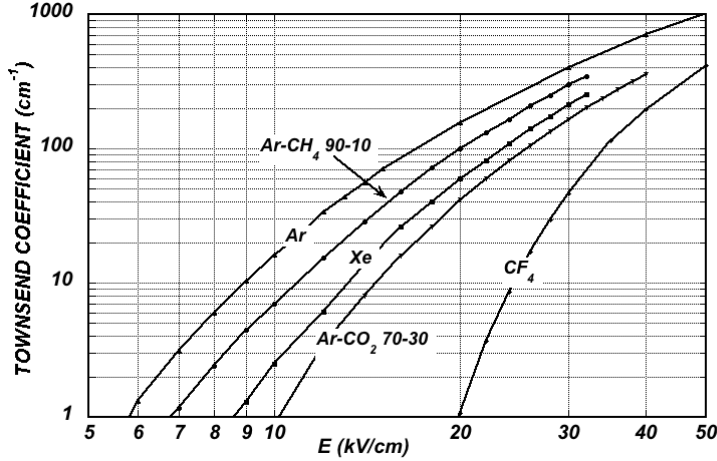
In mixtures containing electronegative molecules, such as  $O_2$  or  $H_2O$ , electrons can be captured to form negative ions. Capture cross-sections are strongly energy-dependent, and therefore the capture probability is a function of applied field. For example, the electron is attached to the oxygen molecule at energies below 1 eV. The three-body electron attachment coefficients may differ greatly for the same additive in different mixtures. As an example, at moderate fields (up to 1 kV/cm) the addition of 0.1% of oxygen to an Ar/ $CO_2$  mixture results in an electron capture probability about twenty times larger than the same addition to Ar/ $CH_4$ .

Carbon tetrafluoride is not electronegative at low and moderate fields, making its use attractive as drift gas due to its very low diffusion. However,  $CF_4$  has a large electron capture cross section at fields above  $\sim 8$  kV/cm, before reaching avalanche field strengths. Depending on detector geometry, some signal reduction and resolution loss can be expected using this gas.

If the electric field is increased sufficiently, electrons gain enough energy between collisions to ionize molecules. Above a gas-dependent threshold, the mean free path for ionization,  $\lambda_i$ , decreases exponentially with the field; its inverse,  $\alpha = 1/\lambda_i$ , is the first Townsend coefficient. In wire chambers, most of the increase of avalanche particle density occurs very close to the anode wires, and a simple electrostatic consideration shows that the largest fraction of the detected signal is due to the motion of positive ions receding from the wires. The electron component, although very fast, contributes very little to the signal. This determines the characteristic shape of the detected signals in the proportional mode: a fast rise followed by a gradual increase. The slow component, the so-called “ion tail” that limits the time resolution of the detector, is usually removed by differentiation of the signal. In uniform fields,  $N_0$  initial electrons multiply over a length  $x$  forming an electron avalanche of size  $N = N_0 e^{\alpha x}$ ;  $N/N_0$  is the gain of the detector. Fig. 1.6 shows examples of Townsend coefficients for several gas mixtures, computed with MAGBOLTZ [65].



**Figure 1.5:** Electron longitudinal diffusion ( $\sigma_L$ ) (dashed lines) and transverse diffusion ( $\sigma_T$ ) (full lines) for 1 cm of drift at NTP and  $B = 0$ . The dotted line shows  $\sigma_T$  for the P10 mixture at 4 T [65].



**Figure 1.6:** Computed first Townsend coefficient  $\alpha$  as a function of electric field in several gases at NTP [65].

Positive ions released by the primary ionization or produced in the avalanches drift and diffuse under the influence of the electric field. Negative ions may also be produced by electron attachment to gas molecules. The drift velocity of ions in the fields encountered in gaseous detectors (up to few kV/cm) is typically about three orders of magnitude less than for electrons. The ion mobility  $\mu$ , the ratio of drift velocity to electric field, is constant for a given ion type up to very high fields. Values of mobility at NTP for ions in their own and other gases are given in Table 1.6 [66]. For different temperatures and pressures, the mobility can be scaled inversely with the density assuming an ideal gas law. For mixtures, due to a very effective charge transfer mechanism, only ions with the lowest ionization potential survive after a short path in the gas. Both the lateral and transverse diffusion of ions are proportional to the square root of the drift time, with a coefficient that depends on temperature but not on the ion mass. Accumulation of ions in the gas drift volume may induce field distortions (see Sec. 1.6.5).

**Table 1.6:** Mobility of ions in gases at NTP [66].

Gas	Ion	Mobility $\mu$ ( $\text{cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ )
He	$\text{He}^+$	10.4
Ne	$\text{Ne}^+$	4.7
Ar	$\text{Ar}^+$	1.54
Ar/ $\text{CH}_4$	$\text{CH}_4^+$	1.87
Ar/ $\text{CO}_2$	$\text{CO}_2^+$	1.72
$\text{CH}_4$	$\text{CH}_4^+$	2.26
$\text{CO}_2$	$\text{CO}_2^+$	1.09

## 24 1. Detectors at accelerators

**1.6.2. Multi-Wire Proportional and Drift Chambers:** Revised March 2010 by Fabio Sauli (CERN) and Maxim Titov (CEA Saclay).

Single-wire counters that detect the ionization produced in a gas by a charged particle, followed by charge multiplication and collection around a thin wire have been used for decades. Good energy resolution is obtained in the proportional amplification mode, while very large saturated pulses can be detected in the streamer and Geiger modes [3].

Multiwire proportional chambers (MWPCs) [67,68], introduced in the late '60's, detect, localize and measure energy deposit by charged particles over large areas. A mesh of parallel anode wires at a suitable potential, inserted between two cathodes, acts almost as a set of independent proportional counters (see Fig. 1.7a). Electrons released in the gas volume drift towards the anodes and produce avalanches in the increasing field. Analytic expressions for the electric field can be found in many textbooks. The fields close to the wires  $E(r)$ , in the drift region  $E_D$ , and the capacitance  $C$  per unit length of anode wire are approximately given by

$$E(r) = \frac{CV_0}{2\pi\epsilon_0} \frac{1}{r} \quad E_D = \frac{CV_0}{2\epsilon_0 s} \quad C = \frac{2\pi\epsilon_0}{\pi(\ell/s) - \ln(2\pi a/s)} \quad (1.14)$$

where  $r$  is the distance from the center of the anode,  $s$  the wire spacing,  $\ell$  and  $V_0$  the distance and potential difference between anode and cathode, and  $a$  the anode wire radius.

Because of electrostatic forces, anode wires are in equilibrium only for a perfect geometry. Small deviations result in forces displacing the wires alternatively below and above the symmetry plane, sometimes with catastrophic results. These displacement forces are countered by the mechanical tension of the wire, up to a maximum unsupported stable length,  $L_M$  [58], above which the wire deforms:

$$L_M = \frac{s}{CV_0} \sqrt{4\pi\epsilon_0 T_M} \quad (1.15)$$

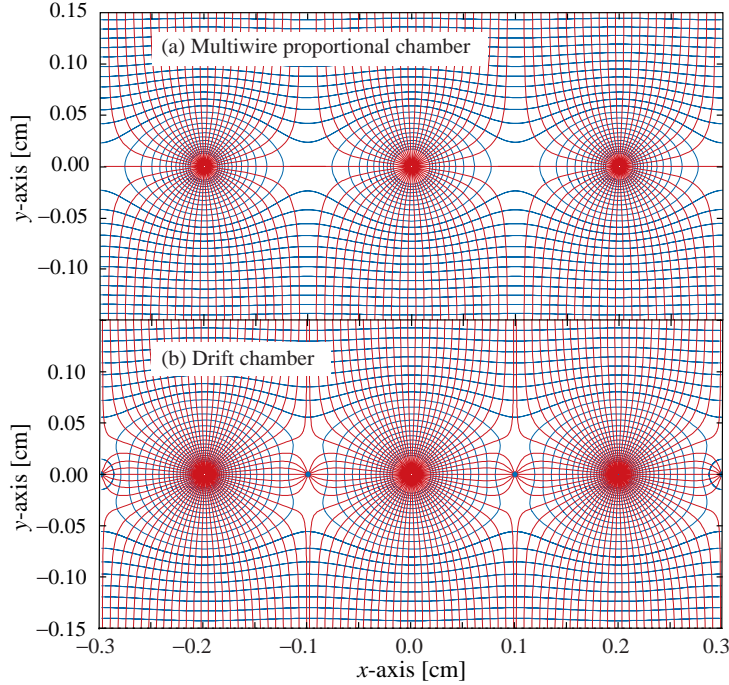
The maximum tension  $T_M$  depends on the wire diameter and modulus of elasticity. Table 1.7 gives approximate values for tungsten and the corresponding maximum stable wire length under reasonable assumptions for the operating voltage ( $V_0 = 5$  kV) [69]. Internal supports and spacers can be used in the construction of longer detectors to overcome limits on the wire length imposed by Eq. (1.15).

**Table 1.7:** Maximum tension  $T_M$  and stable unsupported length  $L_M$  for tungsten wires with spacing  $s$ , operated at  $V_0 = 5$  kV. No safety factor is included.

Wire diameter ( $\mu\text{m}$ )	$T_M$ (newton)	$s$ (mm)	$L_M$ (cm)
10	0.16	1	25
20	0.65	2	85

Detection of charge on the wires over a predefined threshold provides the transverse coordinate to the wire with an accuracy comparable to that of the wire spacing. The coordinate along each wire can be obtained by measuring the ratio of collected charge at the two ends of resistive wires. Making use of the charge profile induced on segmented cathodes, the so-called center-of gravity (COG) method, permits localization of tracks to sub-mm accuracy. Due to the statistics of energy loss and asymmetric ionization clusters, the





**Figure 1.7:** Electric field lines and equipotentials in (a) a multiwire proportional chamber and (b) a drift chamber.

position accuracy is  $\sim 50 \mu\text{m}$  rms for tracks perpendicular to the wire plane, but degrades to  $\sim 250 \mu\text{m}$  at  $30^\circ$  to the normal [70]. The intrinsic bi-dimensional characteristic of the COG readout has found numerous applications in medical imaging.

Drift chambers, developed in the early '70's, can be used to estimate the longitudinal position of a track by exploiting the arrival time of electrons at the anodes if the time of interaction is known [71]. The distance between anode wires is usually several cm, allowing coverage of large areas at reduced cost. In the original design, a thicker wire (the field wire) at the proper voltage, placed between the anode wires, reduces the field at the mid-point between anodes and improves charge collection (Fig. 1.7b). In some drift chamber designs, and with the help of suitable voltages applied to field-shaping electrodes, the electric field structure is adjusted to improve the linearity of space-to-drift-time relation, resulting in better spatial resolution [72].

Drift chambers can reach a longitudinal spatial resolution from timing measurement of order  $100 \mu\text{m}$  (rms) or better for minimum ionizing particles, depending on the geometry and operating conditions. However, a degradation of resolution is observed [73] due to primary ionization statistics for tracks close to the anode wires, caused by the spread in arrival time of the nearest ionization clusters. The effect can be reduced by operating the detector at higher pressures. Sampling the drift time on rows of anodes led to the concept of multiple arrays such as the multi-drift module [74] and the JET chamber [75]. A measurement of drift time, together with the recording of charge sharing from the two ends of the anode wires provides the coordinates of segments of tracks. The total charge gives information on the differential energy loss and is exploited for particle identification. The time projection chamber (TPC) [76] combines a measurement of drift time and charge induction on cathodes, to obtain excellent tracking

for high multiplicity topologies occurring at moderate rates (see Sec. 1.6.5). In all cases, a good knowledge of electron drift velocity and diffusion properties is required. This has to be combined with the knowledge of the electric fields in the structures, computed with commercial or custom-developed software [65,77]. For an overview of detectors exploiting the drift time for coordinate measurement see Refs. 6 and 58.

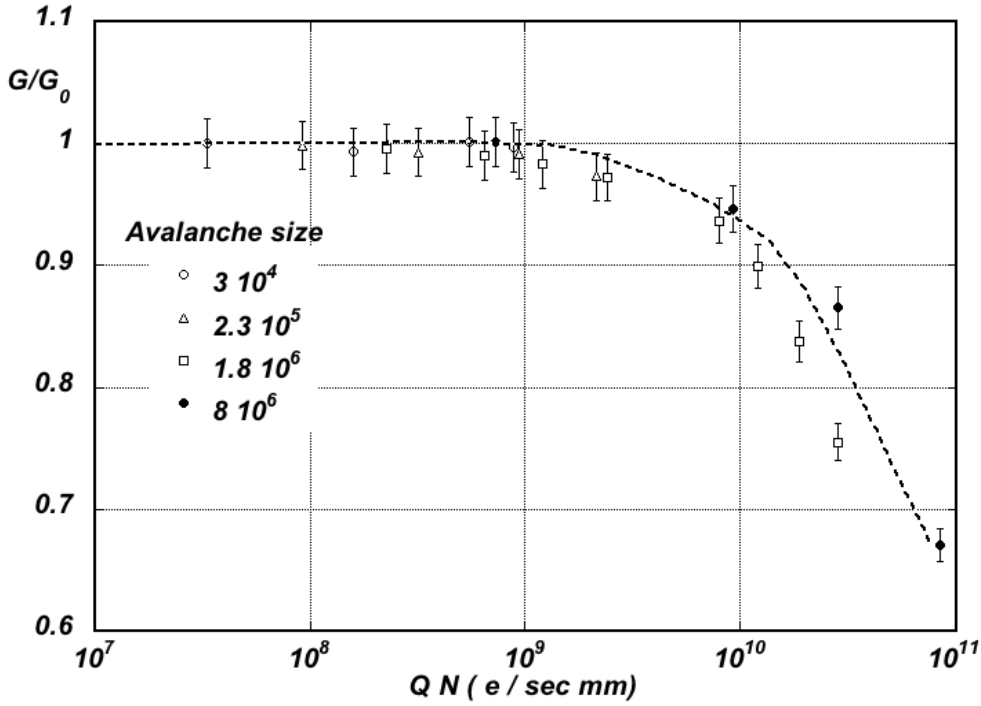
Multiwire and drift chambers have been operated with a variety of gas fillings and operating modes, depending on experimental requirements. The so-called “Magic Gas,” a mixture of argon, isobutane and Freon [68], permits very high and saturated gains ( $\sim 10^6$ ). This gas mixture was used in early wire chambers, but was found to be susceptible to severe aging processes. With present-day electronics, proportional gains around  $10^4$  are sufficient for detection of minimum ionizing particles, and noble gases with moderate amounts of polyatomic gases, such as methane or carbon dioxide, are used.

Although very powerful in terms of performance, multi-wire structures have reliability problems when used in harsh or hard-to-access environments, since a single broken wire can disable the entire detector. Introduced in the '80's, straw and drift tube systems make use of large arrays of wire counters encased in individual enclosures, each acting as an independent wire counter [78]. Techniques for low-cost mass production of these detectors have been developed for large experiments, such as the Transition Radiation Tracker and the Drift Tubes arrays for CERN's LHC experiments [79].

**1.6.3. *High Rate Effects:*** Revised March 2010 by Fabio Sauli (CERN) and Maxim Titov (CEA Saclay).

The production of positive ions in the avalanches and their slow drift before neutralization result in a rate-dependent accumulation of positive charge in the detector. This may result in significant field distortion, gain reduction and degradation of spatial resolution. As shown in Fig. 1.8 [80], the proportional gain drops above a charge production rate around  $10^9$  electrons per second and mm of wire, independently of the avalanche size. For a proportional gain of  $10^4$  and 100 electrons per track, this corresponds to a particle flux of  $10^3 \text{ s}^{-1} \text{ mm}^{-1}$  ( $1 \text{ kHz/mm}^2$  for 1 mm wire spacing).

At high radiation fluxes, a fast degradation of detectors due to the formation of polymers deposits (aging) is often observed. The process has been extensively investigated, often with conflicting results. Several causes have been identified, including organic pollutants and silicone oils. Addition of small amounts of water in many (but not all) cases has been shown to extend the lifetime of the detectors. Addition of fluorinated gases (*e.g.*,  $\text{CF}_4$ ) or oxygen may result in an etching action that can overcome polymer formation, or even eliminate already existing deposits. However, the issue of long-term survival of gas detectors with these gases is controversial [81]. Under optimum operating conditions, a total collected charge of a few coulombs per cm of wire can usually be reached before noticeable degradation occurs. This corresponds, for one mm spacing and at a gain of  $10^4$ , to a total particle flux of  $\sim 10^{14}$  MIPs/cm<sup>2</sup>.



**Figure 1.8:** Charge rate dependence of normalized gas gain  $G/G_0$  (relative to zero counting rate) in proportional thin-wire detectors [80].  $Q$  is the total charge in single avalanche;  $N$  is the particle rate per wire length.

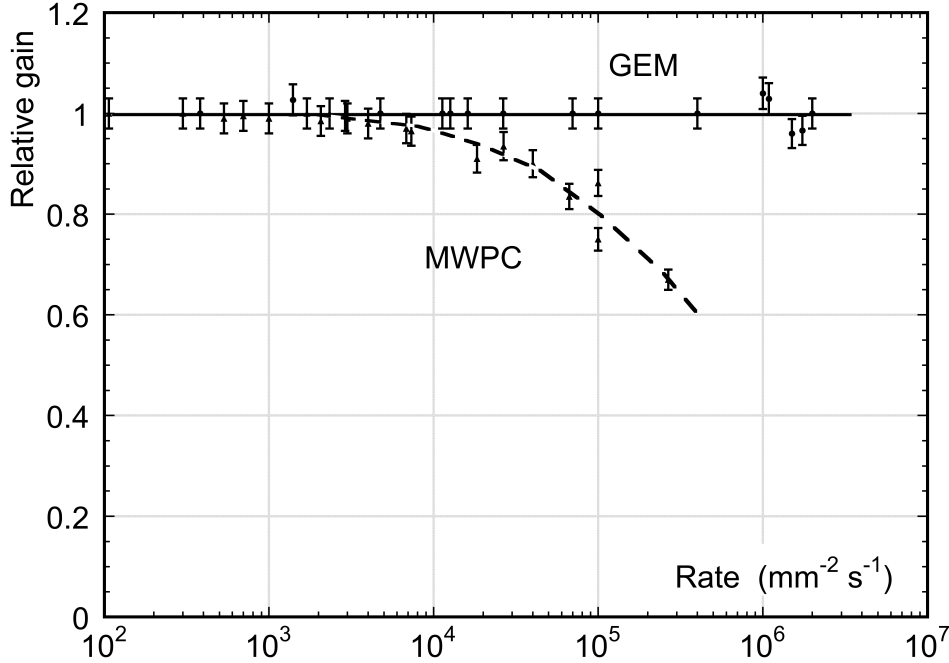
**1.6.4. Micro-Pattern Gas Detectors:** Revised March 2010 by Fabio Sauli (CERN) and Maxim Titov (CEA Saclay)

Despite various improvements, position-sensitive detectors based on wire structures are limited by basic diffusion processes and space charge effects to localization accuracies of  $50\text{--}100\,\mu\text{m}$  [82]. Modern photolithographic technology led to the development of novel Micro-Pattern Gas Detector (MPGD) concepts [83], revolutionizing cell size limitations for many gas detector applications. By using pitch size of a few hundred  $\mu\text{m}$ , an order of magnitude improvement in granularity over wire chambers, these detectors offer intrinsic high rate capability ( $> 10^6\,\text{Hz/mm}^2$ ), excellent spatial resolution ( $\sim 30\,\mu\text{m}$ ), multi-particle resolution ( $\sim 500\,\mu\text{m}$ ), and single photo-electron time resolution in the ns range.

The Micro-Strip Gas Chamber (MSGC), invented in 1988, was the first of the micro-structure gas chambers [84]. It consists of a set of tiny parallel metal strips laid on a thin resistive support, alternatively connected as anodes and cathodes. Owing to the small anode-to-cathode distance ( $\sim 100\,\mu\text{m}$ ), the fast collection of positive ions reduces space charge build-up, and provides a greatly increased rate capability. Unfortunately, the fragile electrode structure of the MSGC turned out to be easily destroyed by discharges induced by heavily ionizing particles [85]. Nevertheless, detailed studies of their properties, and in particular, on the radiation-induced processes leading to discharge breakdown, led to the development of the more powerful devices: GEM and Micromegas. These have improved reliability and radiation hardness. The absence of space-charge effects

## 28 1. Detectors at accelerators

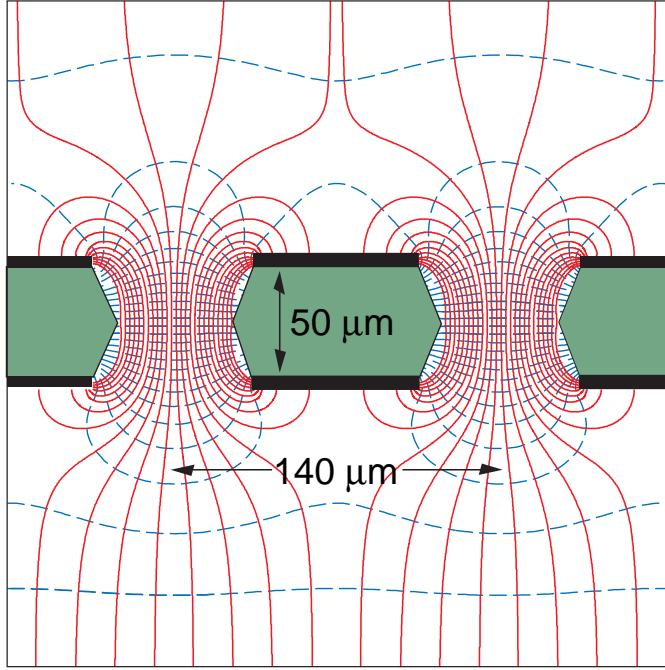
in GEM detectors at the highest rates reached so far and the fine granularity of MPGDs improve the maximum rate capability by more than two orders of magnitude (Fig. 1.9) [72,86]. Even larger rate capability has been reported for Micromegas [87].



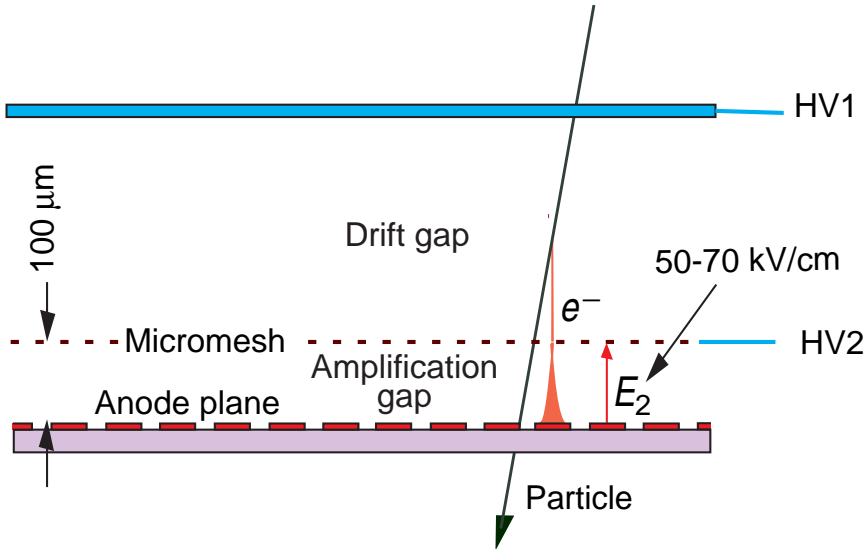
**Figure 1.9:** Normalized gas gain as a function of particle rate for MWPC [72] and GEM [86].

The Gas Electron Multiplier (GEM) detector consists of a thin-foil copper-insulator-copper sandwich chemically perforated to obtain a high density of holes in which avalanches occur [88]. The hole diameter is typically between  $25 \mu\text{m}$  and  $150 \mu\text{m}$ , while the corresponding distance between holes varies between  $50 \mu\text{m}$  and  $200 \mu\text{m}$ . The central insulator is usually (in the original design) the polymer Kapton, with a thickness of  $50 \mu\text{m}$ . Application of a potential difference between the two sides of the GEM generates the electric fields indicated in Fig. 1.10. Each hole acts as an independent proportional counter. Electrons released by the primary ionization particle in the upper conversion region (above the GEM foil) drift into the holes, where charge multiplication occurs in the high electric field ( $50\text{--}70 \text{ kV/cm}$ ). Most of avalanche electrons are transferred into the gap below the GEM. Several GEM foils can be cascaded, allowing the multi-layer GEM detectors to operate at overall gas gain above  $10^4$  in the presence of highly ionizing particles, while strongly reducing the risk of discharges. This is a major advantage of the GEM technology [89]. Localization can then be performed by collecting the charge on a patterned one- or two-dimensional readout board of arbitrary pattern, placed below the last GEM.

The micro-mesh gaseous structure (Micromegas) is a thin parallel-plate avalanche counter, as shown in Fig. 1.11 [90]. It consists of a drift region and a narrow multiplication gap ( $25\text{--}150 \mu\text{m}$ ) between a thin metal grid (micromesh) and the readout electrode (strips or pads of conductor printed on an insulator board). Electrons from



**Figure 1.10:** Schematic view and typical dimensions of the hole structure in the GEM amplification cell. Electric field lines (solid) and equipotentials (dashed) are shown.



**Figure 1.11:** Schematic drawing of the Micromegas detector.

the primary ionization drift through the holes of the mesh into the narrow multiplication gap, where they are amplified. The electric field is homogeneous both in the drift (electric field  $\sim 1$  kV/cm) and amplification (50–70 kV/cm) gaps. In the narrow multiplication region, gain variations due to small variations of the amplification gap are approximately compensated by an inverse variation of the amplification coefficient, resulting in a more uniform gain. The small

## 30 1. *Detectors at accelerators*

amplification gap produces a narrow avalanche, giving rise to excellent spatial resolution:  $12\ \mu\text{m}$  accuracy, limited by the micro-mesh pitch, has been achieved for MIPs, as well as very good time resolution and energy resolution ( $\sim 12\%$  FWHM with 6 keV x rays) [91].

The performance and robustness of GEM and Micromegas have encouraged their use in high-energy and nuclear physics, UV and visible photon detection, astroparticle and neutrino physics, neutron detection and medical physics. Most structures were originally optimized for high-rate particle tracking in nuclear and high-energy physics experiments. COMPASS, a high-luminosity experiment at CERN, pioneered the use of large-area ( $\sim 40 \times 40\ \text{cm}^2$ ) GEM and Micromegas detectors close to the beam line with particle rates of  $25\ \text{kHz}/\text{mm}^2$ . Both technologies achieved a tracking efficiency of close to 100% at gas gains of about  $10^4$ , a spatial resolution of  $70\text{--}100\ \mu\text{m}$  and a time resolution of  $\sim 10\ \text{ns}$ . GEM detectors are also used for triggering in the LHCb Muon System and for tracking in the TOTEM Telescopes. Both GEM and Micromegas devices are foreseen for the upgrade of the LHC experiments and for one of the readout options for the Time Projection Chamber (TPC) at the International Linear Collider (ILC). The development of new fabrication techniques—“bulk” Micromegas technology [92] and single-mask GEMs [93]—is a big step toward industrial production of large-size MPGDs. In some applications requiring very large-area coverage with moderate spatial resolution, coarse macro-patterned detectors, such as Thick GEMs (THGEM) [94] or patterned resistive-plate devices [95] might offer economically interesting solutions.

Sensitive and low-noise electronics enlarge the range of the MPGD applications. Recently, the GEM and Micromegas detectors were read out by high-granularity ( $\sim 50\ \mu\text{m}$  pitch) CMOS chips assembled directly below the GEM or Micromegas amplification structures [96]. These detectors use the bump-bonding pads of a pixel chip as an integrated charge collecting anode. With this arrangement signals are induced at the input gate of a charge-sensitive preamplifier (top metal layer of the CMOS chip). Every pixel is then directly connected to the amplification and digitization circuits, integrated in the underlying active layers of the CMOS technology, yielding timing and charge measurements as well as precise spatial information in 3D.

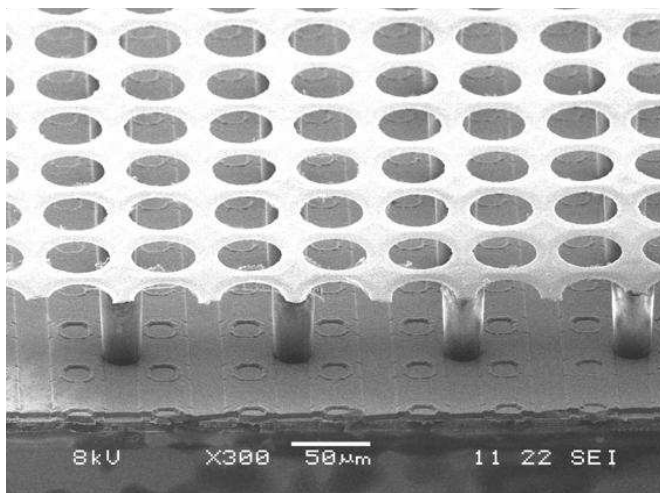
The operation of a MPGD with a Timepix CMOS chip has demonstrated the possibility of reconstructing 3D-space points of individual primary electron clusters with  $\sim 30\ \mu\text{m}$  spatial resolution and event-time resolution with nanosecond precision. This has become indispensable for tracking and triggering and also for discriminating between ionizing tracks and photon conversions. The GEM, in conjunction with a CMOS ASIC,\* can directly view the absorption process of a few keV x-ray quanta and simultaneously reconstruct the direction of emission, which is sensitive to the x-ray polarization. Thanks to these developments, a micro-pattern device with finely segmented CMOS readout can serve as a high-precision “electronic bubble chamber.” This may open new opportunities for x-ray polarimeters, detection of weakly interacting massive particles (WIMPs) and axions, Compton telescopes, and 3D imaging of nuclear recoils.

An elegant solution for the construction of the Micromegas with pixel readout is the integration of the amplification grid and CMOS chip by means of an advanced “wafer post-processing” technology [97].

---

\* Application Specific Integrated Circuit

This novel concept is called “Ingrid” (see Fig. 1.12). With this technique, the structure of a thin ( $1\text{ }\mu\text{m}$ ) aluminum grid is fabricated on top of an array of insulating pillars, which stands  $\sim 50\text{ }\mu\text{m}$  above the CMOS chip. The sub- $\mu\text{m}$  precision of the grid dimensions and avalanche gap size results in a uniform gas gain. The grid hole size, pitch and pattern can be easily adapted to match the geometry of any pixel readout chip.



**Figure 1.12:** Photo of the Micromegas “Ingrid” detector. The grid holes can be accurately aligned with readout pixels of CMOS chip. The insulating pillars are centered between the grid holes, thus avoiding dead regions.

Recent developments in radiation hardness research with state-of-the-art MPGDs are reviewed in Ref. 98. Earlier aging studies of GEM and Micromegas concepts revealed that they might be even less vulnerable to radiation-induced performance degradation than standard silicon microstrip detectors.

The RD51 collaboration was established in 2008 to further advance technological developments of micro-pattern detectors and associated electronic-readout systems for applications in basic and applied research [99].

**1.6.5. Time-projection chambers:** Reviser October 2011 by D. Karlen (U. of Victoria and TRIUMF, Canada)

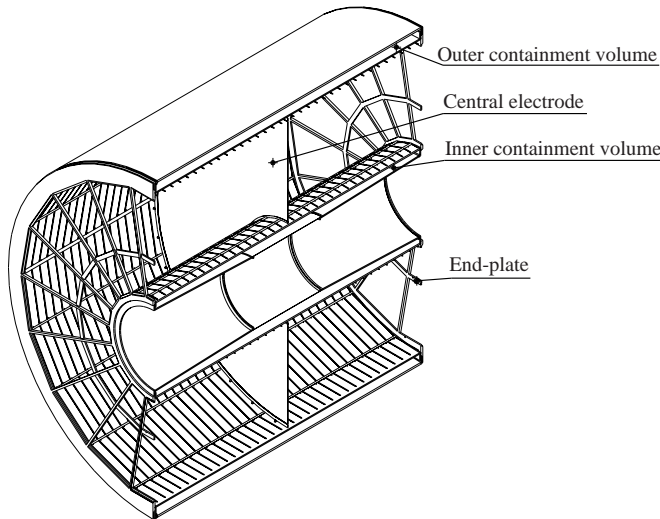
The Time Projection Chamber (TPC) concept, invented by David Nygren in the late 1970’s [76], is the basis for charged particle tracking in a large number of particle and nuclear physics experiments. A uniform electric field drifts tracks of electrons produced by charged particles traversing a medium, either gas or liquid, towards a surface segmented into 2D readout pads. The signal amplitudes and arrival times are recorded to provide full 3D measurements of the particle trajectories. The intrinsic 3D segmentation gives the TPC a distinct advantage over other large volume tracking detector designs which record information only in a 2D projection with less overall segmentation, particularly for pattern recognition in events with large numbers of particles.

## 32 1. Detectors at accelerators

Gaseous TPC's are often designed to operate within a strong magnetic field (typically parallel to the drift field) so that particle momenta can be estimated from the track curvature. For this application, precise spatial measurements in the plane transverse to the magnetic field are most important. Since the amount of ionization along the length of the track depends on the velocity of the particle, ionization and momentum measurements can be combined to identify the types of particles observed in the TPC. The estimator for the energy deposit by a particle is usually formed as the truncated mean of the energy deposits, using the 50%–70% of the samples with the smallest signals. Variance due to energetic  $\delta$ -ray production is thus reduced.

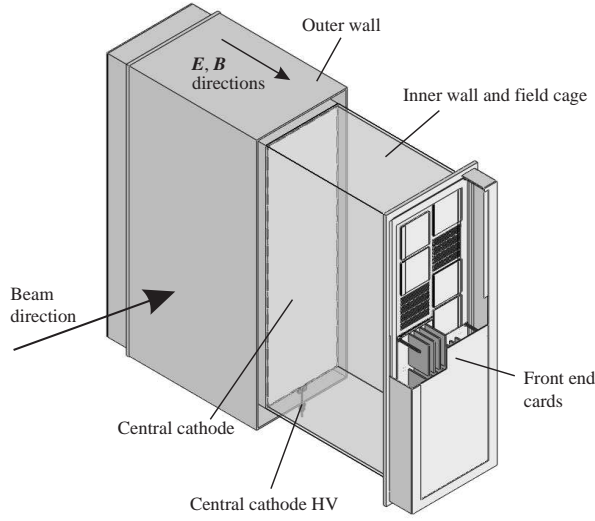
Gas amplification of  $10^3$ – $10^4$  at the readout endplate is usually required in order to provide signals with sufficient amplitude for conventional electronics to sense the drifted ionization. Until recently, the gas amplification system used in TPC's have exclusively been planes of anode wires operated in proportional mode placed close to the readout pads. Performance has been recently improved by replacing these wire planes with micro-pattern gas detectors, namely GEM [88] and Micromegas [90] devices. Advances in electronics miniaturization have been important in this development, allowing pad areas to be reduced to the  $10 \text{ mm}^2$  scale or less, well matched to the narrow extent of signals produced with micro-pattern gas detectors. Presently, the ultimate in fine segmentation TPC readout are silicon sensors, with  $0.05 \text{ mm} \times 0.05 \text{ mm}$  pixels, in combination with GEM or Micromegas [100]. With such fine granularity it is possible to count the number of ionization clusters along the length of a track which, in principle, can improve the particle identification capability.

Examples of two modern large volume gaseous TPC's are shown in Fig. 1.13 and Fig. 1.14. The particle identification performance is illustrated in Fig. 1.15, for the original TPC in the PEP-4/9 experiment [101].

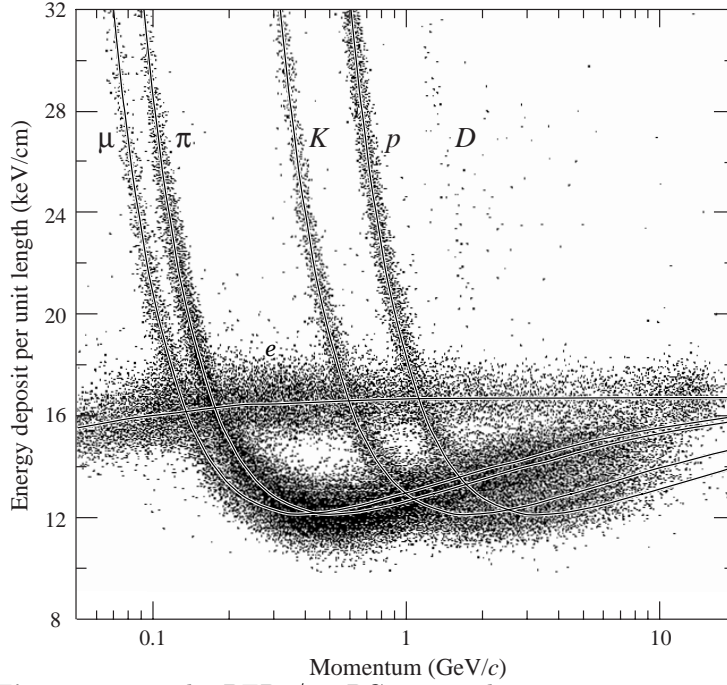


**Figure 1.13:** The ALICE TPC shown in a cutaway view [102]. The drift volume is 5 m long with a 5 m diameter. Gas amplification is provided by planes of anode wires.





**Figure 1.14:** One of the 3 TPC modules for the near detector of the T2K experiment [103]. The drift volume is  $2\text{ m} \times 2\text{ m} \times 0.8\text{ m}$ . Micromegas devices are used for gas amplification and readout.



**Figure 1.15:** The PEP4/9-TPC energy deposit measurements (185 samples, 8.5 atm Ar-CH<sub>4</sub> 80:20). The ionization rate at the Fermi plateau (at high  $\beta$ ) is 1.4 times that for the minimum at lower  $\beta$ . This ratio increases to 1.6 at atmospheric pressure.

The greatest challenges for a large TPC arise from the long drift distance, typically 100 times further than in a comparable wire chamber design. In particular, the long drift distance can make the device sensitive to small distortions in the electric field. Distortions

### 34 1. Detectors at accelerators

can arise from a number of sources, such as imperfections in the TPC construction, deformations of the readout surface, or the presence of ions in the active medium.

For a gaseous TPC operated in a magnetic field, the electron drift velocity  $v$  is defined by Eq. (1.13). With a strong magnetic field parallel to the electric field and a gas with a large value of  $\omega\tau$  (also favored to reduce transverse diffusion as discussed below), the transverse displacements of the drifting electrons due to electric field distortions are reduced. In this mode of operation, it is essential to precisely map the magnetic field as the electron drift lines closely follow the magnetic field lines. Corrections for electric and/or magnetic field non-uniformities can be determined from control samples of electrons produced by ionizing the gas with UV laser beams, from photoelectrons produced on the cathode, or from tracks emanating from calibration reactions.

The long drift distance means that there is a delay, typically 10–100  $\mu\text{s}$  in a large gaseous TPC, for signals to arrive at the endplate. For experiments with shorter intervals between events, this can produce ambiguities in the starting time for the drift of ionization. This can be resolved by matching the TPC data with that from an auxiliary detector providing additional spatial or timing information.

In a gaseous TPC, the motion of positive ions is much slower than the electrons, and so the positive ions produced by many events may exist in the active volume. Of greatest concern is the ions produced in the gas amplification stage. Large gaseous TPC's built until now with wire planes have included a gating grid that prevent the positive ions from escaping into the drift volume in the interval between event triggers. Micro-pattern gas detectors release much less positive ions than wire planes operating at the same gain, which may allow operation of a TPC without a gating grid.

Given the long drift distance in a large TPC, the active medium must remain very pure, as small amounts of contamination can absorb the ionization signal. For example, in a typical large gaseous TPC,  $\text{O}_2$  must be kept below a few parts in  $10^5$ , otherwise a large fraction of the drifting electrons will become attached. Special attention must be made in the choice of construction materials in order to avoid the release of other electronegative contaminants.

Diffusion degrades the position information of ionization that drifts a long distance. For a gaseous TPC, the effect can be alleviated by the choice of a gas with low intrinsic diffusion or by operating in a strong magnetic field parallel to the drift field with a gas which exhibits a significant reduction in transverse diffusion with magnetic field. For typical operation without magnetic field, the transverse extent of the electrons,  $\sigma_{Dx}$ , is a few mm after drifting 1 m due to diffusion. With a strong magnetic field,  $\sigma_{Dx}$  can be reduced by as much as a factor of 10,

$$\sigma_{Dx}(B)/\sigma_{Dx}(0) = \frac{1}{\sqrt{1 + \omega^2\tau^2}} \quad (1.17)$$

where  $\omega\tau$  is defined above. The diffusion limited position resolution from the information collected by a single row of pads is

$$\sigma_x = \frac{\sigma_{Dx}}{\sqrt{n}} \quad (1.18)$$

where  $n$  is the effective number of electrons collected by the pad row, giving an ultimate single row resolution of order 100  $\mu\text{m}$ .

Diffusion is significantly reduced in a negative-ion TPC [104], which uses a special gas mixture that attaches electrons immediately as they

are produced. The drifting negative ions exhibit much less diffusion than electrons. The slow drift velocity and small  $\omega\tau$  of negative ions must be compatible with the experimental environment.

The spatial resolution achieved by a TPC is determined by a number of factors in addition to diffusion. Non-uniform ionization along the length of the track is a particularly important factor, and is responsible for the so-called “track angle” and “ $\mathbf{E} \times \mathbf{B}$ ” effects. If the boundaries between pads in a row are not parallel to the track, the ionization fluctuations will increase the variance in the position estimate from that row. For this reason, experiments with a preferred track direction should have pad boundaries aligned with that direction. Traditional TPC’s with wire plane amplification suffer from the effects of non-parallel electric and magnetic fields near the wires that rotate ionization segments, thereby degrading the resolution because of the non-uniform ionization. Micro-pattern gas detectors exhibit a much smaller  $\mathbf{E} \times \mathbf{B}$  effect, since their feature size is much smaller than that of a wire grid.

**1.6.6. Transition radiation detectors (TRD’s):** Written August 2007 by P. Nevski (BNL) and A. Romaniouk (Moscow Eng. & Phys. Inst.)

Transition radiation (TR) x rays are produced when a highly relativistic particle ( $\gamma \gtrsim 10^3$ ) crosses a refractive index interface, as discussed in Sec. 30.7. The x rays, ranging from a few keV to a few dozen keV, are emitted at a characteristic angle  $1/\gamma$  from the particle trajectory. Since the TR yield is about 1% per boundary crossing, radiation from multiple surface crossings is used in practical detectors. In the simplest concept, a detector module might consist of low- $Z$  foils followed by a high- $Z$  active layer made of proportional counters filled with a Xe-rich gas mixture. The atomic number considerations follow from the dominant photoelectric absorption cross section per atom going roughly as  $Z^n/E_x^3$ , where  $n$  varies between 4 and 5 over the region of interest, and the x-ray energy is  $E_x$ .<sup>\*</sup> To minimize self-absorption, materials such as polypropylene, Mylar, carbon, and (rarely) lithium are used as radiators. The TR signal in the active regions is in most cases superimposed upon the particle’s ionization losses. These drop a little faster than  $Z/A$  with increasing  $Z$ , providing another reason for active layers with high  $Z$ .

The TR intensity for a single boundary crossing always increases with  $\gamma$ , but for multiple boundary crossings interference leads to saturation near a Lorentz factor  $\gamma_{\text{sat}} = 0.6 \omega_1 \sqrt{\ell_1 \ell_2} / c$  [105], where  $\omega_1$  is the radiator plasma frequency,  $\ell_1$  is its thickness, and  $\ell_2$  the spacing. In most of the detectors used in particle physics the radiator parameters are chosen to provide  $\gamma_{\text{sat}} \approx 2000$ . Those detectors normally work as threshold devices, ensuring the best electron/pion separation in the momentum range  $1 \text{ GeV}/c \lesssim p \lesssim 150 \text{ GeV}/c$ .

One can distinguish two design concepts—“thick” and “thin” detectors:

1. The radiator, optimized for a minimum total radiation length at maximum TR yield and total TR absorption, consists of few hundred foils (for instance 300 20  $\mu\text{m}$  thick polypropylene foils). A dominant fraction of the soft TR photons is absorbed in the radiator itself. To increase the average TR photon energy further,

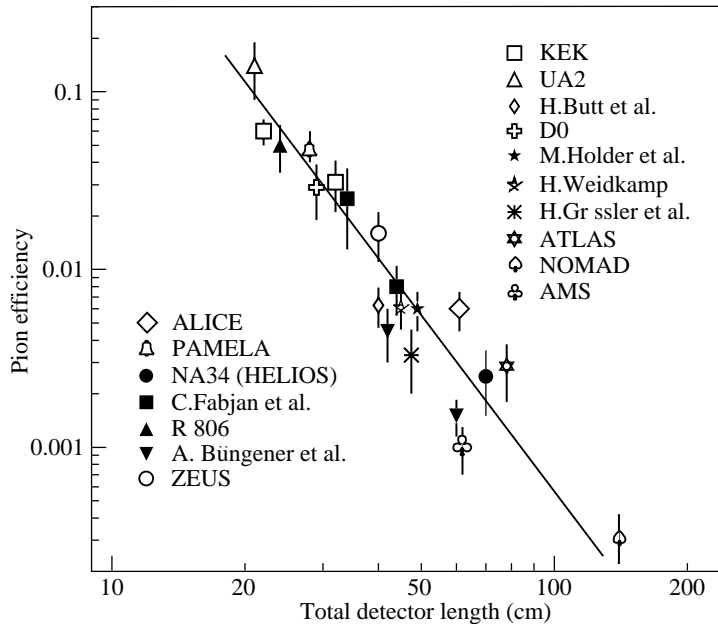
---

<sup>\*</sup> Photon absorption coefficients for the elements (via a NIST link), and  $dE/dx|_{\text{min}}$  and plasma energies for many materials are given in [pdg.lbl.gov/AtomicNuclearProperties](http://pdg.lbl.gov/AtomicNuclearProperties).

part of the radiator far from the active layers is often made of thicker foils. The detector thickness, about 2 cm for Xe-filled gas chambers, is optimized to absorb the shaped x-ray spectrum. A classical detector is composed of several similar modules which respond nearly independently. Such detectors were used in the NA34 [106], and are being used in the ALICE experiment [107].

2. In another TRD concept a fine granular radiator/detector structure exploits the soft part of the TR spectrum more efficiently. This can be achieved, for instance, by distributing small-diameter straw-tube detectors uniformly or in thin layers throughout the radiator material (foils or fibers). Even with a relatively thin radiator stack, radiation below 5 keV is mostly lost in the radiators themselves. However for photon energies above this value the absorption becomes smaller and the radiation can be registered by several consecutive detector layers, thus creating a strong TR build-up effect. Descriptions of detectors using this approach can be found in both accelerator and space experiments [107]. For example, in the ATLAS TR tracker charged particles cross about 35 effective straw tube layers embedded in the radiator material [107]. The effective thickness of the Xe gas per straw is about 2.3 mm and the average number of foils per straw is about 40 with an effective foil thickness of about 20  $\mu\text{m}$ .

Both TR photon absorption and the TR build-up significantly affect the detector performance. Although the values mentioned above are typical for most of the plastic radiators used with Xe-based detectors, they vary significantly depending on detector parameters: radiator material, thickness and spacing, the construction of the sensitive chambers, their position, *etc.* Thus careful simulations are usually needed to build a detector optimized for a particular application.



**Figure 1.16:** Pion efficiency measured (or predicted) for different TRDs as a function of the detector length for a fixed electron efficiency of 90%. The plot is taken from [106] with efficiencies of more recent detectors added [107].

The discrimination between electrons and pions can be based on the charge deposition measured in each detection module, on the number of clusters—energy depositions observed above an optimal threshold (usually in the 5 to 7 keV region), or on more sophisticated methods analyzing the pulse shape as a function of time. The total energy measurement technique is more suitable for thick gas volumes, which absorb most of the TR radiation and where the ionization loss fluctuations are small. The cluster-counting method works better for detectors with thin gas layers, where the fluctuations of the ionization losses are big. Cluster-counting replaces the Landau-Vavilov distribution of background ionization energy losses with the Poisson statistics of  $\delta$ -electrons, responsible for the distribution tails. The latter distribution is narrower than the Landau-Vavilov distribution.

The major factor in the performance of any TRD is its overall length. This is illustrated in Fig. 1.16, which shows, for a variety of detectors, the pion efficiency at a fixed electron efficiency of 90% as a function of the overall detector length. The experimental data, covering a range of particle energies from a few GeV to 40 GeV, are rescaled to an energy of 10 GeV when possible. Phenomenologically, the rejection power against pions increases as  $5 \cdot 10^{L/38}$ , where the range of validity is  $L \approx 20\text{--}100$  cm.

Many recent TRDs combine particle identification with charged-track measurement in the same detector [107]. This provides a powerful tool for electron identification even at very high particle densities. Another example of this combination is described by Brigida *et al.* in Ref. 107. In this work Si-microstrip detectors operating in a magnetic field are used both for particle and TR detection. The excellent coordinate resolution of the Si detectors allows spatial separation of the TR photons from particle ionization tracks with relatively modest distances between radiator and detector.

Recent TRDs for particle astrophysics are designed to directly measure the Lorentz factor of high-energy nuclei by using the quadratic dependence of the TR yield on nuclear charge; see Cherry and Müller papers in Ref. 107. The radiator configuration  $(\ell_1, \ell_2)$  is tuned to extend the TR yield rise up to  $\gamma \lesssim 10^5$  using more energetic part of the TR spectrum (up to 100 keV). Exotic radiator materials such as aluminum and unusual TR detection methods (Compton scattering) are used in such cases.

**1.6.7. Resistive-plate chambers:** Revised September 2007 by H.R. Band (U. Wisconsin).

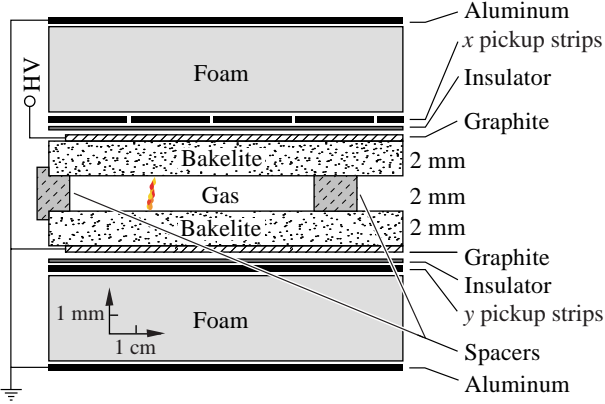
The resistive-plate chamber (RPC) was developed by Santonico and Cardarelli in the early 1980's [108] as a low-cost alternative to large scintillator planes.\* Most commonly, an RPC is constructed from two parallel high-resistivity ( $10^9\text{--}10^{13}$   $\Omega\text{-cm}$ ) glass or phenolic (Bakelite)/melamine laminate plates with a few-mm gap between them which is filled with atmospheric-pressure gas. The gas is chosen to absorb UV photons in order to limit transverse growth of discharges. The backs of the plates are coated with a lower-resistivity paint or ink ( $\sim 10^5$   $\Omega/\square$ ), and a high potential (7–12 kV) is maintained between them. The passage of a charged particle initiates an electric discharge, whose size and duration are limited since the current reduces the local potential to below that needed to maintain the discharge. The sensitivity of the detector outside of this region is

---

\* It was based on earlier work on a spark counter with one high-resistivity plate [109].

### 38 1. Detectors at accelerators

unaffected. The signal readout is via capacitive coupling to metallic strips on both sides of the detector which are separated from the high voltage coatings by thin insulating sheets. The  $x$  and  $y$  position of the discharge can be measured if the strips on opposite sides of the gap are orthogonal. When operated in streamer mode, the induced signals on the strips can be quite large ( $\sim 300$  mV), making sensitive electronics unnecessary. An example of an RPC structure is shown in Fig. 1.17.



**Figure 1.17:** Schematic cross section of a typical RPC, in this case the single-gap streamer-mode BaBar RPC.

RPC's have inherent rate limitations since the time needed to re-establish the field after a discharge is proportional to the chamber capacitance and plate resistance. The average charge per streamer is 100–1000 pC. Typically, the efficiency of streamer-mode glass RPC's begins to fall above  $\sim 0.4$  Hz/cm<sup>2</sup>. Because of Bakelite's lower bulk resistivity, Bakelite RPC's can be efficient at 10–100 Hz/cm<sup>2</sup>. The need for higher rate capability led to the development of avalanche-mode RPC's, in which the gas and high voltage have been tuned to limit the growth of the electric discharge, preventing streamer formation. Typical avalanche-mode RPC's have a signal charge of about 10 pC and can be efficient at 1 kHz/cm<sup>2</sup>. The avalanche discharge produces a much smaller induced signal on the pickup strips ( $\sim 1$  mV) than streamers, and thus requires a more sophisticated and careful electronic design.

Many variations of the initial RPC design have been built for operation in either mode. Efficiencies of  $\gtrsim 92\%$  for single gaps can be improved by the use of two or more gas gaps with shared pickup strips. Non-flammable and more environmentally friendly gas mixtures have been developed. In streamer mode, various mixtures of argon with isobutane and tetrafluoroethane have been used. For avalanche mode operation, a gas mixture of tetrafluoroethane (C<sub>2</sub>H<sub>2</sub>F<sub>4</sub>) with 2–5% isobutane and 0.4–10% sulfur hexafluoride (SF<sub>6</sub>) is typical. An example of large-scale RPC use is provided by the muon system being built for the ATLAS detector, where three layers of pairs of RPC's are used to trigger the drift tube arrays between the pairs. The total area is about 10,000 m<sup>2</sup>. These RPC's provide a spatial resolution of 1 cm and a time resolution of 1 ns at an efficiency  $\geq 99\%$ .

Developments of multiple-gap RPC's [110] lead to RPC designs with much better timing resolution ( $\sim 50$  ps) for use in time-of-flight particle identification systems. A pioneering design used by the HARP

experiment [111] has two sets of 2 thin gas gaps (0.3 mm) separated by thin (0.7 mm) glass plates. The outer plates are connected to high voltage and ground while the inner plate is electrically isolated and floats to a stable equilibrium potential. The observed RPC intrinsic time resolution of 127 ps may have been limited by amplifier noise. Fonte provides useful review [112] of other RPC designs.

Operational experience with RPC's has been mixed. Several experiments (*e.g.*, L3 and HARP) have reported reliable performance. However, the severe problems experienced with the BaBar RPC's have raised concerns about the long-term reliability of Bakelite RPC's.

Glass RPC's have had fewer problems, as seen by the history of the BELLE chambers. A rapid growth in the noise rate and leakage current in some of the BELLE glass RPC's was observed during commissioning. It was found that water vapor in the input gas was reacting with fluorine (produced by the disassociation of the tetrafluoroethane in the streamers) to produce hydrofluoric acid. The acid etched the glass surfaces, leading to increased noise rates and lower efficiencies. The use of copper gas piping to insure the dryness of the input gas stopped the problem. The BELLE RPC's have now operated reliably for more than 5 years.

Several different failure modes diagnosed in the first-generation BaBar Bakelite RPC's caused the average efficiency of the barrel RPC's to fall from  $\gtrsim 90\%$  to 35% in five years. The linseed oil which is used in Bakelite RPC's to coat the inner surface [113] had not been completely cured. Under warm conditions (32°C) and high voltage, oil collected on the spacers between the gaps or formed oil-drop bridges between the gaps. This led to large leakage currents (50–100  $\mu\text{A}$  in some chambers) which persisted even when the temperature was regulated at 20°C. In addition, the graphite layer used to distribute the high voltage over the Bakelite became highly resistive (100 k $\Omega/\square \rightarrow$  10 M $\Omega/\square$ ), resulting in lowered efficiency in some regions and the complete death of whole chambers.

The BaBar problems and the proposed use of Bakelite RPC's in the LHC detectors prompted detailed studies of RPC aging and have led to improved construction techniques and a better understanding of RPC operational limits. The graphite layer has been improved and should be stable with integrated currents of  $\lesssim 600 \text{ mC/cm}^2$ . Molded gas inlets and improved cleanliness during construction have reduced the noise rate of new chambers. Unlike glass RPC's, Bakelite RPC's have been found to require humid input gases to prevent drying of the Bakelite (increasing the bulk resistivity) which would decrease the rate capability. Second-generation BaBar RPC's incorporating many of the above improvements have performed reliably for over two years [114].

With many of these problems solved, new-generation RPC's are now being or soon will be used in about a dozen cosmic-ray and HEP detectors. Their comparatively low cost, ease of construction, good time resolution, high efficiency, and moderate spatial resolution make them attractive in many situations, particularly those requiring fast timing and/or large-area coverage.

## 40 1. Detectors at accelerators

### 1.7. Semiconductor detectors

Updated August 2011 by H. Spieler.

#### 1.7.1. Materials Requirements:

Semiconductor detectors are essentially solid state ionization chambers. Absorbed energy forms electron-hole pairs, *i.e.*, negative and positive charge carriers, which under an applied electric field move towards their respective collection electrodes, where they induce a signal current. The energy required to form an electron-hole pair is proportional to the bandgap. In tracking detectors the energy loss in the detector should be minimal, whereas for energy spectroscopy the stopping power should be maximized, so for gamma rays high- $Z$  materials are desirable.

Measurements on silicon photodiodes [116] show that for photon energies below 4 eV one electron-hole ( $e-h$ ) pair is formed per incident photon. The mean energy  $E_i$  required to produce an  $e-h$  pair peaks at 4.4 eV for a photon energy around 6 eV. Above  $\sim 1.5$  keV it assumes a constant value, 3.67 eV at room temperature. It is larger than the bandgap energy because momentum conservation requires excitation of lattice vibrations (phonons). For minimum-ionizing particles, the most probable charge deposition in a 300  $\mu\text{m}$  thick silicon detector is about 3.5 fC (22000 electrons). Other typical ionization energies are 2.96 eV in Ge, 4.2 eV in GaAs, and 4.43 eV in CdTe.

Since both electronic and lattice excitations are involved, the variance in the number of charge carriers  $N = E/E_i$  produced by an absorbed energy  $E$  is reduced by the Fano factor  $F$  (about 0.1 in Si and Ge). Thus,  $\sigma_N = \sqrt{FN}$  and the energy resolution  $\sigma_E/E = \sqrt{FE_i/E}$ . However, the measured signal fluctuations are usually dominated by electronic noise or energy loss fluctuations in the detector. The electronic noise contributions depend on the pulse shaping in the signal processing electronics, so the choice of the shaping time is critical (see Sec. 1.8).

A smaller bandgap would produce a larger signal and improve energy resolution, but the intrinsic resistance of the material is critical. Thermal excitation, given by the Fermi-Dirac distribution, promotes electrons into the conduction band, so the thermally excited carrier concentration increases exponentially with decreasing bandgaps. In pure Si the carrier concentration is  $\sim 10^{10} \text{ cm}^{-3}$  at 300 K, corresponding to a resistivity  $\rho \approx 400 \text{ k}\Omega \text{ cm}$ . In reality, crystal imperfections and minute impurity concentrations limit Si carrier concentrations to  $\sim 10^{11} \text{ cm}^{-3}$  at 300 K, corresponding to a resistivity  $\rho \approx 40 \text{ k}\Omega \text{ cm}$ . In practice, resistivities up to 20  $\text{k}\Omega \text{ cm}$  are available, with mass production ranging from 5 to 10  $\text{k}\Omega \text{ cm}$ . Signal currents at keV scale energies are of order  $\mu\text{A}$ . However, for a resistivity of  $10^4 \text{ }\Omega \text{ cm}$  a 300  $\mu\text{m}$  thick sensor with 1  $\text{cm}^2$  area would have a resistance of 300  $\Omega$ , so 30 V would lead to a current flow of 100 mA and a power dissipation of 3 W. On the other hand, high-quality single crystals of Si and Ge can be grown economically with suitably large volumes, so to mitigate the effect of resistivity one resorts to reverse-biased diode structures. Although this reduces the bias current relative to a resistive material, the thermally excited leakage current can still be excessive at room temperature, so Ge diodes are typically operated at liquid nitrogen temperature (77 K).

A major effort is to find high- $Z$  materials with a bandgap that is sufficiently high to allow room-temperature operation while still providing good energy resolution. Compound semiconductors, *e.g.*, CdZnTe, can allow this, but typically suffer from charge collection



problems, characterized by the product  $\mu\tau$  of mobility and carrier lifetime. In Si and Ge  $\mu\tau > 1 \text{ cm}^2 \text{ V}^{-1}$  for both electrons and holes, whereas in compound semiconductors it is in the range  $10^{-3}$ – $10^{-8}$ . Since for holes  $\mu\tau$  is typically an order of magnitude smaller than for electrons, detector configurations where the electron contribution to the charge signal dominates—*e.g.*, strip or pixel structures—can provide better performance.

### 1.7.2. Detector Configurations:

A  $p$ - $n$  junction operated at reverse bias forms a sensitive region depleted of mobile charge and sets up an electric field that sweeps charge liberated by radiation to the electrodes. Detectors typically use an asymmetric structure, *e.g.*, a highly doped  $p$  electrode and a lightly doped  $n$  region, so that the depletion region extends predominantly into the lightly doped volume.

In a planar device the thickness of the depleted region is

$$W = \sqrt{2\epsilon (V + V_{bi}) / Ne} = \sqrt{2\rho\mu\epsilon (V + V_{bi})}, \quad (1.19)$$

where  $V$  = external bias voltage

$V_{bi}$  = “built-in” voltage ( $\approx 0.5 \text{ V}$  for resistivities typically used in Si detectors)

$N$  = doping concentration

$e$  = electronic charge

$\epsilon$  = dielectric constant =  $11.9 \epsilon_0 \approx 1 \text{ pF/cm}$  in Si

$\rho$  = resistivity (typically 1–10  $\text{k}\Omega \text{ cm}$  in Si)

$\mu$  = charge carrier mobility

=  $1350 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$  for electrons in Si

In Si =  $450 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$  for holes in Si

$$W = 0.5 \left[ \frac{\mu\text{m}}{\sqrt{\Omega\text{-cm} \cdot \text{V}}} \right] \times \sqrt{\rho (V + V_{bi})} \text{ for } n\text{-type Si, and}$$

$$W = 0.3 \left[ \frac{\mu\text{m}}{\sqrt{\Omega\text{-cm} \cdot \text{V}}} \right] \times \sqrt{\rho (V + V_{bi})} \text{ for } p\text{-type Si.}$$

The conductive  $p$  and  $n$  regions together with the depleted volume form a capacitor with the capacitance per unit area

$$C = \epsilon / W \approx 1 \text{ [pF/cm]} / W \text{ in Si.} \quad (1.22)$$

In strip and pixel detectors the capacitance is dominated by the fringing capacitance. For example, the strip-to-strip Si fringing capacitance is  $\sim 1$ – $1.5 \text{ pF cm}^{-1}$  of strip length at a strip pitch of 25–50  $\mu\text{m}$ .

Large volume ( $\sim 10^2$ – $10^3 \text{ cm}^3$ ) Ge detectors are commonly configured as coaxial detectors, *e.g.*, a cylindrical  $n$ -type crystal with 5–10 cm diameter and 10 cm length with an inner 5–10 mm diameter  $n^+$  electrode and an outer  $p^+$  layer forming the diode junction. Ge can be grown with very low impurity levels,  $10^9$ – $10^{10} \text{ cm}^{-3}$  (HPGe), so these large volumes can be depleted with several kV.

## 42 1. Detectors at accelerators

### 1.7.3. Signal Formation:

The signal pulse shape depends on the instantaneous carrier velocity  $v(x) = \mu E(x)$  and the electrode geometry, which determines the distribution of induced charge (*e.g.*, see Ref. 115, pp. 71–83). Charge collection time decreases with increasing bias voltage, and can be reduced further by operating the detector with “overbias,” *i.e.*, a bias voltage exceeding the value required to fully deplete the device. The collection time is limited by velocity saturation at high fields (in Si approaching  $10^7$  cm/s at  $E > 10^4$  V/cm); at an average field of  $10^4$  V/cm the collection time is about 15 ps/ $\mu\text{m}$  for electrons and 30 ps/ $\mu\text{m}$  for holes. In typical fully-depleted detectors 300  $\mu\text{m}$  thick, electrons are collected within about 10 ns, and holes within about 25 ns.

Position resolution is limited by transverse diffusion during charge collection (typically 5  $\mu\text{m}$  for 300  $\mu\text{m}$  thickness) and by knock-on electrons. Resolutions of 2–4  $\mu\text{m}$  (rms) have been obtained in beam tests. In magnetic fields, the Lorentz drift deflects the electron and hole trajectories and the detector must be tilted to reduce spatial spreading (see “Hall effect” in semiconductor textbooks).

Electrodes can be in the form of cm-scale pads, strips, or  $\mu\text{m}$ -scale pixels. Various readout structures have been developed for pixels, *e.g.*, CCDs, DEPFETs, monolithic pixel devices that integrate sensor and electronics (MAPS), and hybrid pixel devices that utilize separate sensors and readout ICs connected by two-dimensional arrays of solder bumps. For an overview and further discussion see Ref. 115.

In gamma ray spectroscopy ( $E_\gamma > 10^2$  keV) Compton scattering dominates, so for a significant fraction of events the incident gamma energy is not completely absorbed, *i.e.*, the Compton scattered photon escapes from the detector and the energy deposited by the Compton electron is only a fraction of the total. Distinguishing multi-interaction events, *e.g.*, multiple Compton scatters with a final photoelectric absorption, from single Compton scatters allows background suppression. Since the individual interactions take place in different parts of the detector volume, these events can be distinguished by segmenting the outer electrode of a coaxial detector and analyzing the current pulse shapes. The different collection times can be made more distinguishable by using “point” electrodes, where most of the signal is induced when charges are close to the electrode, similarly to strip or pixel detectors. Charge clusters arriving from different positions in the detector will arrive at different times and produce current pulses whose major components are separated in time. Point electrodes also reduce the electrode capacitance, which reduces electronic noise, but careful design is necessary to avoid low-field regions in the detector volume.

The increase in reverse bias current due to bulk damage is  $\Delta I_r = \alpha \Phi$  per unit volume, where  $\Phi$  is the particle fluence and  $\alpha$  the damage coefficient ( $\alpha \approx 3 \times 10^{-17}$  A/cm for minimum ionizing protons and pions after long-term annealing;  $\alpha \approx 2 \times 10^{-17}$  A/cm for 1 MeV neutrons). The reverse bias current depends strongly on temperature

$$\frac{I_R(T_2)}{I_R(T_1)} = \left(\frac{T_2}{T_1}\right)^2 \exp \left[ -\frac{E}{2k} \left( \frac{T_1 - T_2}{T_1 T_2} \right) \right], \quad (1.23)$$

where  $E = 1.2$  eV, so rather modest cooling can reduce the current substantially ( $\sim 6$ -fold current reduction in cooling from room temperature to 0°C).

Displacement damage forms acceptor-like states. These trap electrons, building up a negative space charge, which in turn requires

an increase in the applied voltage to sweep signal charge through the detector thickness. This has the same effect as a change in resistivity, *i.e.*, the required voltage drops initially with fluence, until the positive and negative space charge balance and very little voltage is required to collect all signal charge. At larger fluences the negative space charge dominates, and the required operating voltage increases ( $V \propto N$ ). The safe limit on operating voltage ultimately limits the detector lifetime. Strip detectors specifically designed for high voltages have been extensively operated at bias voltages  $>500$  V. Since the effect of radiation damage depends on the electronic activity of defects, various techniques have been applied to neutralize the damage sites. For example, additional doping with oxygen can increase the allowable charged hadron fluence roughly three-fold [117]. Detectors with columnar electrodes normal to the surface can also extend operational lifetime [118]. The increase in leakage current with fluence, on the other hand, appears to be unaffected by resistivity and whether the material is *n* or *p*-type. At fluences beyond  $10^{15}$  cm $^{-2}$  decreased carrier lifetime becomes critical [119,120].

Strip and pixel detectors have remained functional at fluences beyond  $10^{15}$  cm $^{-2}$  for minimum ionizing protons. At this damage level, charge loss due to recombination and trapping becomes significant and the high signal-to-noise ratio obtainable with low-capacitance pixel structures extends detector lifetime. The higher mobility of electrons makes them less sensitive to carrier lifetime than holes, so detector configurations that emphasize the electron contribution to the charge signal are advantageous, *e.g.*,  $n^+$  strips or pixels on a *p*- or *n*-substrate. The occupancy of the defect charge states is strongly temperature dependent; competing processes can increase or decrease the required operating voltage. It is critical to choose the operating temperature judiciously ( $-10$  to  $0^\circ\text{C}$  in typical collider detectors) and limit warm-up periods during maintenance. For a more detailed summary see Ref. 121 and the web-sites of the ROSE and RD50 collaborations at <http://RD48.web.cern.ch/rd48> and <http://RD50.web.cern.ch/rd50>. Materials engineering, *e.g.*, introducing oxygen interstitials, can improve certain aspects and is under investigation. At high fluences diamond is an alternative, but operates as an insulator rather than a reverse-biased diode.

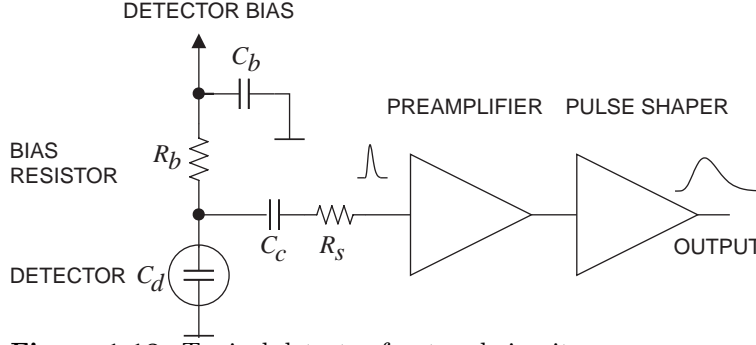
Currently, the lifetime of detector systems is still limited by the detectors; in the electronics use of standard “deep submicron” CMOS fabrication processes with appropriately designed circuitry has increased the radiation resistance to fluences  $> 10^{15}$  cm $^{-2}$  of minimum ionizing protons or pions. For a comprehensive discussion of radiation effects see Ref. 122.

## 1.8. Low-noise electronics

Revised August 2011 by H. Spieler.

Many detectors rely critically on low-noise electronics, either to improve energy resolution or to allow a low detection threshold. A typical detector front-end is shown in Fig. 1.18.

The detector is represented by a capacitance  $C_d$ , a relevant model for most detectors. Bias voltage is applied through resistor  $R_b$  and the signal is coupled to the preamplifier through a blocking capacitor  $C_c$ . The series resistance  $R_s$  represents the sum of all resistances present in the input signal path, *e.g.* the electrode resistance, any input protection networks, and parasitic resistances in the input transistor. The preamplifier provides gain and feeds a pulse shaper, which tailors the overall frequency response to optimize signal-to-noise ratio while



**Figure 1.18:** Typical detector front-end circuit.

limiting the duration of the signal pulse to accommodate the signal pulse rate. Even if not explicitly stated, all amplifiers provide some form of pulse shaping due to their limited frequency response.

Electronic noise originates in the detector, in the series resistance, in the bias resistor, and the amplifier. The first contributes a noise current with the spectral density  $dP_n/df \propto i_n^2 = 2eI_d$  (“shot noise”), the second acts as a noise voltage generator with the spectral density  $e_n^2 = 4kTR_s$ , and the third acts as a noise current generator,  $i_{nb}^2 = 4kT/R_b$ . Here  $e$  is the electronic charge,  $k$  the Boltzmann constant, and  $T$  the temperature. The amplifier noise is described by the voltage and current noise densities  $e_{na}$  and  $i_{na}$ . Frequently, amplifiers (and other components) exhibit additional low frequency (“ $1/f$ ”) noise, characterized by the spectral density  $e_{nf}^2 = A_f/f$ .

Since radiation detectors typically convert the deposited energy into charge, the system’s noise level is conveniently expressed as an equivalent noise charge  $Q_n$ , which is equal to the detector signal that yields a signal-to-noise ratio of one. For the circuit shown above,

$$Q_n^2 = \left( 2eI_d + 4kT/R_b + i_{na}^2 \right) F_i T_S + (4kTR_s + e_{na}^2) F_v C_d^2 / T_S + F_{vf} A_f C_d^2. \quad (1.30)$$

where  $C$  is the sum of all capacitances shunting the input,  $F_i$ ,  $F_v$ , and  $F_{vf}$  depend on the shape of the pulse determined by the shaper and  $T_S$  is a characteristic time, for example the peaking time of a semi-gaussian pulse or the sampling interval in a correlated double sample.

For quick estimates, one can use the following equation, which assumes an FET amplifier (negligible  $i_{na}$ ) and a simple  $CR$ - $RC$  shaper with time constants  $\tau$  (equal to the peaking time):

$$(Q_n/e)^2 = 12 \left[ \frac{1}{\text{nA} \cdot \text{ns}} \right] I_d \tau + 6 \times 10^5 \left[ \frac{\text{k}\Omega}{\text{ns}} \right] \frac{\tau}{R_b} + 3.6 \times 10^4 \left[ \frac{\text{ns}}{(\text{pF})^2 (\text{nV})^2 / \text{Hz}} \right] e_n^2 \frac{C^2}{\tau}. \quad (1.31)$$

In bipolar transistors the base current adds additional shot noise, which can be included in the bias current term.

In timing measurements, the slope-to-noise ratio must be optimized, rather than the signal-to-noise ratio alone, so the rise time  $t_r$  of the pulse is important. The “jitter”  $\sigma_t$  of the timing distribution is

$$\sigma_t = \frac{\sigma_n}{(dS/dt)_{S_T}} \approx \frac{t_r}{S/N}, \quad (1.34)$$

where  $\sigma_n$  is the rms noise and the derivative of the signal  $dS/dt$  is evaluated at the trigger level  $S_T$ . To increase  $dS/dt$  without incurring excessive noise, the amplifier bandwidth should match the rise-time of the detector signal. The 10 to 90% rise time of an amplifier with bandwidth  $f_U$  is  $0.35/f_U$ . For example, an oscilloscope with 350 MHz bandwidth has a 1 ns rise time. When amplifiers are cascaded, which is invariably necessary, the individual rise times add in quadrature.

$$t_r \approx \sqrt{t_{r1}^2 + t_{r2}^2 + \dots + t_{rn}^2} . \quad (1.35)$$

Increasing signal-to-noise ratio also improves time resolution, so minimizing the total capacitance at the input is also important. At high signal-to-noise ratios, the time jitter can be much smaller than the rise time. The timing distribution may shift with signal level (“walk”), but this can be corrected by various means, either in hardware or software [8].

The basic principles discussed above apply to both analog and digital signal processing. In digital signal processing the pulse shaper shown in Fig. 1.18 is replaced by an analog to digital converter (ADC) followed by a digital processor that determines the pulse shape. Digital signal processing allows great flexibility in implementing filtering functions. The software can be changed readily to adapt to a wide variety of operating conditions and it is possible to implement filters that are impractical or even impossible using analog circuitry. However, this comes at the expense of increased circuit complexity and increased demands on the ADC compared to analog shaping.

If the sampling rate of the ADC is too low, high frequency components will be transferred to lower frequencies (“aliasing”). The sampling rate of the ADC must be high enough to capture the maximum frequency component of the input signal. Apart from missing information on the fast components of the pulse, undersampling introduces spurious artifacts. If the frequency range of the input signal is much greater, the noise at the higher frequencies will be transferred to lower frequencies and increase the noise level in the frequency range of pulses formed in the subsequent digital shaper. The Nyquist criterion states that the sampling frequency must be at least twice the maximum relevant input frequency. This requires that the bandwidth of the circuitry preceding the ADC must be limited. The most reliable technique is to insert a low-pass filter.

The digitization process also introduces inherent noise, since the voltage range  $\Delta V$  corresponding to a minimum bit introduces quasi-random fluctuations relative to the exact amplitude

$$\sigma_n = \frac{\Delta V}{\sqrt{12}} . \quad (1.36)$$

When the Nyquist condition is fulfilled the noise bandwidth  $\Delta f_n$  is spread nearly uniformly and extends to 1/2 the sampling frequency  $f_S$ , so the spectral noise density

$$e_n = \frac{\sigma_n}{\sqrt{\Delta f_n}} = \frac{\Delta V}{\sqrt{12}} \cdot \frac{1}{\sqrt{f_S/2}} = \frac{\Delta V}{\sqrt{6f_S}} . \quad (1.37)$$

Sampling at a higher frequency spreads the total noise over a larger frequency range, so oversampling can be used to increase the effective resolution. In practice, this quantization noise is increased by differential nonlinearity. Furthermore, the equivalent input noise of ADCs is often rather high, so the overall gain of the stages preceding the ADC must be sufficiently large for the preamplifier input noise to override.

When implemented properly, digital signal processing provides significant advantages in systems where the shape of detector signal pulses changes greatly, for example in large semiconductor detectors for gamma rays or in gaseous detectors (*e.g.* TPCs) where the duration of the current pulse varies with drift time, which can range over orders of magnitude. Where is analog signal processing best (most efficient)? In systems that require fast time response the high power requirements of high-speed ADCs are prohibitive. Systems that are not sensitive to pulse shape can use fixed shaper constants and rather simple filters, which can be either continuous or sampled. In high density systems that require small circuit area and low power (*e.g.* strip and pixel detectors), analog filtering often yields the required response and tends to be most efficient.

For a more detailed introduction to detector signal processing and electronics see Ref. 115.

### 1.9. Calorimeters

A calorimeter is designed to measure the energy deposition and its direction for a contained electromagnetic (EM) or hadronic shower. The characteristic interaction distance for an electromagnetic interaction is the radiation length  $X_0$ , which ranges from  $13.8 \text{ g cm}^{-2}$  in iron to  $6.0 \text{ g cm}^{-2}$  in uranium.\* Similarly, the characteristic nuclear interaction length  $\lambda_I$  varies from  $132.1 \text{ g cm}^{-2}$  (Fe) to  $209 \text{ g cm}^{-2}$  (U).† In either case, a calorimeter must be many interaction lengths deep, where “many” is determined by physical size, cost, and other factors. EM calorimeters tend to be  $15\text{--}30 X_0$  deep, while hadronic calorimeters are usually compromised at  $5\text{--}8 \lambda_I$ . In real experiments there is likely to be an EM calorimeter in front of the hadronic section, which in turn has less sampling density in the back, so the hadronic cascade occurs in a succession of different structures.

In all cases there is a premium on small  $\lambda_I/\rho$  and  $X_0/\rho$  (both with units of length). These quantities are shown for  $Z > 20$  for the chemical elements in Fig. 1.21. For the hadronic case, metallic absorbers in the W–Au region are best, followed by U. The Ru–Pd region elements are rare and expensive. Lead is a bad choice. Given cost considerations, Fe and Cu might be appropriate choices. For EM calorimeters high  $Z$  is preferred, and lead is not a bad choice.

These considerations are for *sampling calorimeters* consisting of metallic absorber sandwiched or (threaded) with an active material which generates signal. The active medium may be a scintillator, an ionizing noble liquid, a gas chamber, a semiconductor, or a Cherenkov radiator. The average interaction length is thus greater than that of the absorber alone, sometimes substantially so.

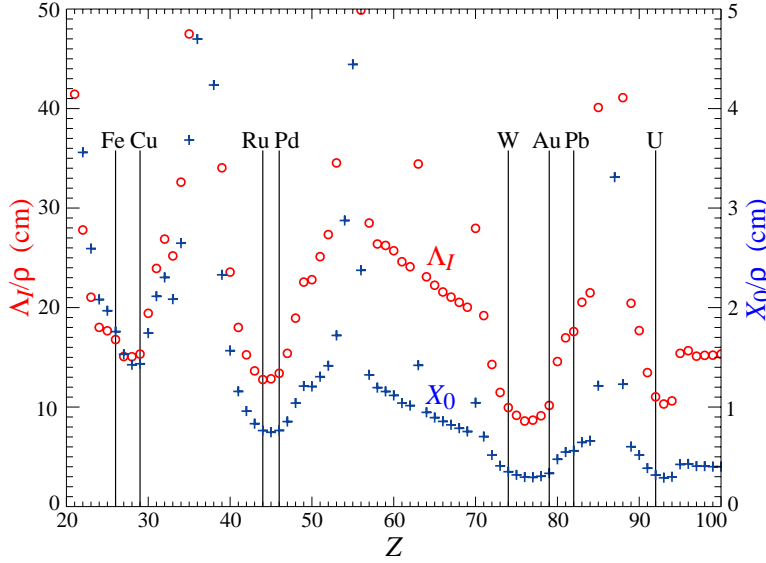
There are also *homogeneous calorimeters*, in which the entire volume is sensitive, *i.e.*, contributes signal. Homogeneous calorimeters (so far usually electromagnetic) may be built with inorganic heavy (high density, high  $\langle Z \rangle$ ) scintillating crystals, or non-scintillating Cherenkov radiators such as lead glass and lead fluoride. Scintillation light and/or ionization in noble liquids can be detected. Nuclear interaction lengths in inorganic crystals range from  $17.8 \text{ cm}$  ( $\text{LuAlO}_3$ ) to  $42.2 \text{ cm}$  ( $\text{NaI}$ ). Popular choices have been BGO with  $\lambda_I = 22.3 \text{ cm}$  and  $X_0 = 1.12 \text{ cm}$ , and  $\text{PbWO}_4$  ( $20.3 \text{ cm}$  and  $0.89 \text{ cm}$ ). Properties of

---

\*  $X_0 = 120 \text{ g cm}^{-2} Z^{-2/3}$  to better than 5% for  $Z > 23$ .

†  $\lambda_I = 37.8 \text{ g cm}^{-2} A^{0.312}$  to within 0.8% for  $Z > 15$ .

See [pdg.lbl.gov/AtomicNuclearProperties](http://pdg.lbl.gov/AtomicNuclearProperties) for actual values.



**Figure 1.21:** Nuclear interaction length  $\lambda_I/\rho$  (circles) and radiation length  $X_0/\rho$  (+’s) in cm for the chemical elements with  $Z > 20$  and  $\lambda_I < 50$  cm.

these and other commonly used inorganic crystal scintillators can be found in Table 1.4.

#### 1.9.1. *Electromagnetic calorimeters:*

Revised October 2009 by R.-Y. Zhu (California Inst. of Technology).

The development of electromagnetic showers is discussed in the section on “Passage of Particles Through Matter” (Sec. 30 of this *Review*).

Formulae are given which approximately describe average showers, but since the physics of electromagnetic showers is well understood, detailed and reliable Monte Carlo simulation is possible. EGS4 [125] and GEANT [126] have emerged as the standards.

There are homogeneous and sampling electromagnetic calorimeters. In a homogeneous calorimeter the entire volume is sensitive, *i.e.*, contributes signal. Homogeneous electromagnetic calorimeters may be built with inorganic heavy (high- $Z$ ) scintillating crystals such as BGO, CsI, NaI, and PWO, non-scintillating Cherenkov radiators such as lead glass and lead fluoride, or ionizing noble liquids. Properties of commonly used inorganic crystal scintillators can be found in Table 1.4. A sampling calorimeter consists of an active medium which generates signal and a passive medium which functions as an absorber. The active medium may be a scintillator, an ionizing noble liquid, a gas chamber, or a semiconductor. The passive medium is usually a material of high density, such as lead, iron, copper, or depleted uranium.

The energy resolution  $\sigma_E/E$  of a calorimeter can be parametrized as  $a/\sqrt{E} \oplus b \oplus c/E$ , where  $\oplus$  represents addition in quadrature and  $E$  is in GeV. The stochastic term  $a$  represents statistics-related fluctuations such as intrinsic shower fluctuations, photoelectron statistics, dead material at the front of the calorimeter, and sampling fluctuations. For a fixed number of radiation lengths, the stochastic term  $a$  for a sampling calorimeter is expected to be proportional to  $\sqrt{t/f}$ , where

$t$  is plate thickness and  $f$  is sampling fraction [127,128]. While  $a$  is at a few percent level for a homogeneous calorimeter, it is typically 10% for sampling calorimeters. The main contributions to the systematic, or constant, term  $b$  are detector non-uniformity and calibration uncertainty. In the case of the hadronic cascades discussed below, non-compensation also contributes to the constant term. One additional contribution to the constant term for calorimeters built for modern high-energy physics experiments, operated in a high-beam intensity environment, is radiation damage of the active medium. This can be minimized by developing radiation-hard active media [48] and by frequent *in situ* calibration and monitoring [47,128]. With effort, the constant term  $b$  can be reduced to below one percent. The term  $c$  is due to electronic noise summed over readout channels within a few Molière radii. The best energy resolution for electromagnetic shower measurement is obtained in total absorption homogeneous calorimeters, *e.g.* calorimeters built with heavy crystal scintillators. These are used when ultimate performance is pursued.

The position resolution depends on the effective Molière radius and the transverse granularity of the calorimeter. Like the energy resolution, it can be factored as  $a/\sqrt{E} \oplus b$ , where  $a$  is a few to 20 mm and  $b$  can be as small as a fraction of mm for a dense calorimeter with fine granularity. Electromagnetic calorimeters may also provide direction measurement for electrons and photons. This is important for photon-related physics when there are uncertainties in event origin, since photons do not leave information in the particle tracking system. Typical photon angular resolution is about  $45 \text{ mrad}/\sqrt{E}$ , which can be provided by implementing longitudinal segmentation [129] for a sampling calorimeter or by adding a preshower detector [130] for a homogeneous calorimeter without longitudinal segmentation.

### 1.9.2. *Hadronic calorimeters:* [1–5,128]

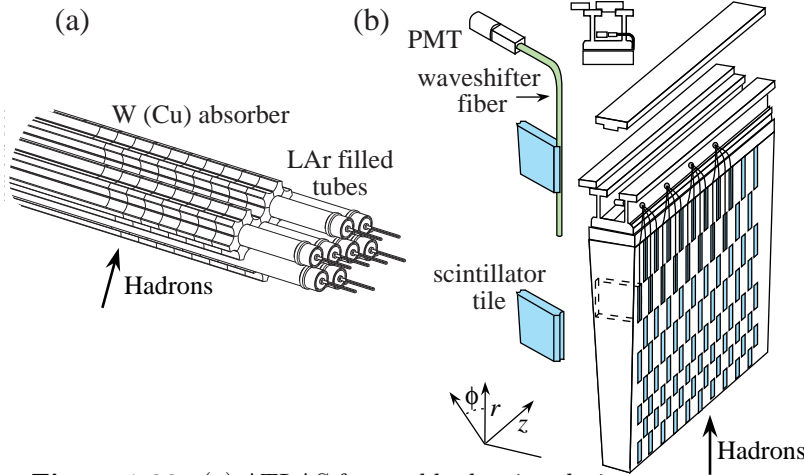
Revised October 2011 by D. E. Groom (LBNL).

Most large hadron calorimeters are parts of large  $4\pi$  detectors at colliding beam facilities. At present these are sampling calorimeters: plates of absorber (Fe, Pb, Cu, or occasionally U or W) alternating with plastic scintillators (plates, tiles, bars), liquid argon (LAr), or gaseous detectors. The ionization is measured directly, as in LAr calorimeters, or via scintillation light observed by photodetectors (usually PMT's). Waveshifting fibers are often used to solve difficult problems of geometry and light collection uniformity. Silicon sensors are being studied for ILC detectors; in this case  $e$ - $h$  pairs are collected. There are as many variants of these schemes as there are calorimeters, including variations in geometry of the absorber and sensors, *e.g.*, scintillating fibers threading an absorber [131], and the “accordion” LAr detector, with zig-zag absorber plates to minimize channeling effects. Another departure from the traditional sandwich structure is the LAr-tube design shown in Fig. 1.22(a).

A relatively new variant is the detection of Cerenkov light in hadron calorimetry. Such a calorimeter is sensitive to  $e^\pm$ 's in the EM showers plus a few relativistic pions. An example is the radiation-hard forward calorimeter in CMS, with iron absorber and quartz fiber readout by PMT's.

Ideally the calorimeter is segmented in  $\phi$  and  $\theta$  (or  $\eta = -\ln \tan(\theta/2)$ ). Fine segmentation, while desirable, is limited by cost, readout complexity, practical geometry, and the transverse size of the cascades—but see [132]. An example, a wedge of the ATLAS central barrel calorimeter, is shown in Fig. 1.22(b).





**Figure 1.22:** (a) ATLAS forward hadronic calorimeter structure (FCal2, 3). Tubes containing LAr are embedded in a mainly tungsten matrix. (b) ATLAS central calorimeter wedge; iron with plastic scintillator tile with wavelength-shifting fiber readout.

In an inelastic hadronic collision a significant fraction  $f_{em}$  of the energy is removed from further hadronic interaction by the production of secondary  $\pi^0$ 's and  $\eta$ 's, whose decay photons generate high-energy electromagnetic (EM) showers. Charged secondaries ( $\pi^\pm$ ,  $p$ , ...) deposit energy via ionization and excitation, but also interact with nuclei, producing spallation protons and neutrons, evaporation neutrons, and spallation products. The charged collision products produce detectable ionization, as do the showering  $\gamma$ -rays from the prompt de-excitation of highly excited nuclei. The recoiling nuclei generate little or no detectable signal. The neutrons lose kinetic energy in elastic collisions over hundreds of ns, gradually thermalize, and are captured, with the production of more  $\gamma$ -rays—usually outside the acceptance gate of the electronics. Between endothermic spallation losses, nuclear recoils, and late neutron capture, a significant fraction of the hadronic energy (20%–40%, depending on the absorber and energy of the incident particle) is invisible.

In contrast to EM showers, hadronic cascade processes are characterized by the production of relatively few high-energy particles. The lost energy and  $f_{em}$ , the  $\pi^0 \rightarrow \gamma\gamma$  fraction are highly variable from event to event. Until there is event-by-event knowledge of both the invisible energy loss and EM deposit (to be discussed below), the energy resolution of a hadron calorimeter will remain significantly worse than that of an EM calorimeter.

It has been shown by a simple induction argument, and verified by experiment, that the decrease in the average value of the hadronic energy fraction ( $\langle f_h \rangle = 1 - \langle f_{em} \rangle$ ) as the projectile energy  $E$  increases is fairly well described by the power law [133,135]

$$\langle f_h \rangle \approx (E/E_0)^{m-1} \quad (\text{for } E > E_0), \quad (1.38)$$

at least up to a few hundred GeV. The exponent  $m$  depends logarithmically on the mean multiplicity and the mean fractional loss to  $\pi^0$  production in a single interaction. It is in the range 0.80–0.87, but must be obtained experimentally for each calorimeter configuration.  $E_0$  is roughly the energy for the onset of inelastic collisions. It is 1 GeV or a little less for incident pions [133].

In a hadron-nucleus collision a large fraction of the incident energy is carried by a “leading particle” with the same quark content as the incident hadron. If the projectile is a charged pion, the leading particle is usually a pion, which can be neutral and hence contributes to the EM sector. This is not true for incident protons. The result is an increased mean hadronic fraction for incident protons: in Eq. (1.39b)  $E_0 \approx 2.6$  GeV [133,134,136].

The EM energy deposit is usually detected more efficiently than the hadronic energy deposit. If the detection efficiency for the EM sector is  $e$  and that for the hadronic sector is  $h$ , then the ratio of the mean response to a pion to that for an electron is

$$\pi/e = \langle f_{em} \rangle + \langle f_h \rangle h/e = 1 - (1 - h/e) \langle f_h \rangle \quad (1.39a)$$

$$\approx 1 - (1 - h/e) (E/E_0)^{m-1} . \quad (1.39b)$$

If  $h \neq e$  the hadronic response is not a linear function of energy. Only the product  $(1 - h/e)E_0^{1-m}$  can be obtained by measuring  $\pi/e$  as a function of energy. Since  $1 - m$  is small and  $E_0 \approx 1$  GeV for the usual pion-induced cascades, this fact is usually ignored and  $h/e$  is reported.

The discussion above assumes an idealized calorimeter, with the same structure throughout and without leakage. “Real” calorimeters usually have an EM detector in front and a coarse “catcher” in the back. Complete containment is generally impractical.

By definition,  $0 \leq f_{em} \leq 1$ . Its variance changes only slowly with energy, but perforce  $\langle f_{em} \rangle \rightarrow 1$  as the projectile energy increases. An empirical power law  $\sigma_{f_{em}} = (E/E_1)^{1-\ell}$  (where  $\ell < 1$ ) describes the energy dependence adequately and has the right asymptotic properties. For  $h/e \neq 1$ , fluctuations in  $f_{em}$  significantly contribute to the resolution, in particular contributing a larger fraction of the variance at high energies. Since the  $f_{em}$  distribution has a tail on the high side, the calorimeter response is non-Gaussian with a high-energy tail if  $h/e < 1$ . *Noncompensation* ( $h/e \neq 1$ ) thus seriously degrades resolution as well as producing a nonlinear response.

It is clearly desirable to *compensate* the response, *i.e.*, to design the calorimeter such that  $h/e = 1$ . This is possible only with a sampling calorimeter, where several variables can be chosen or tuned:

1. Decrease the EM sensitivity. EM cross sections increase with  $Z$ ,<sup>†</sup> and most of the energy in an EM shower is deposited by low-energy electrons. A disproportionate fraction of the EM energy is thus deposited in the higher- $Z$  absorber. Lower- $Z$  cladding, such as the steel cladding on ZEUS U plates, preferentially absorbs low-energy  $\gamma$ 's in EM showers and thus also lowers the electronic response. G10 signal boards in the DØ calorimeters had the same effect. The degree of EM signal suppression can be somewhat controlled by tuning the sensor/absorber thickness ratio.
2. Increase the hadronic sensitivity. The abundant neutrons produced in the cascade have a large  $n$ - $p$  scattering cross section, with the production of low-energy scattered protons in hydrogenous sampling materials such as butane-filled proportional counters or plastic scintillator. (When scattering from a nucleus with mass number  $A$ , a neutron can at most lose  $4/(1+A)^2$  of its kinetic energy.) The down side in the scintillator case is that the signal from a highly-ionizing proton stub can be reduced by as much as 90% by recombination and quenching (Birk's Law, Eq. (1.2)).

---

<sup>†</sup> The asymptotic pair-production cross section scales roughly as  $Z^{0.75}$ , and  $|dE/dx|$  slowly decreases with increasing  $Z$ .

3. Fabjan and Willis proposed that the additional signal generated in the aftermath of fission in  $^{238}\text{U}$  absorber plates should compensate nuclear fluctuations [137]. The production of fission fragments due to fast  $n$  capture was later observed [138]. However, while a very large amount of energy is released, it is mostly carried by low-velocity, very highly ionizing fission fragments which produce very little observable signal because of recombination and quenching. The approach seemed promising for awhile. But, for example, much of the compensation observed with the ZEUS  $^{238}\text{U}$ /scintillator calorimeter was mainly the result of methods 2 and 3 above.

Motivated very much by the work of Brau, Gabriel, Brückmann, and Wigmans [139], several groups built calorimeters which were very nearly compensating. The degree of compensation was sensitive to the acceptance gate width, and so could be somewhat further tuned. These included

- a) HELIOS with 2.5 mm thick scintillator plates sandwiched between 2 mm thick  $^{238}\text{U}$  plates (one of several structures);  $\sigma/E = 0.34/\sqrt{E}$  was obtained,
- b) ZEUS, 2.6 cm thick scintillator plates between 3.3 mm  $^{238}\text{U}$  plates;  $\sigma/E = 0.35/\sqrt{E}$ ,
- c) a ZEUS prototype with 10 mm Pb plates and 2.5 mm scintillator sheets;  $\sigma/E = 0.44/\sqrt{E}$ , and
- d) DØ, where the sandwich cell consists of a 4–6 mm thick  $^{238}\text{U}$  plate, 2.3 mm LAr, a G-10 signal board, and another 2.3 mm LAr gap.

Given geometrical and cost constraints, the calorimeters used in modern collider detectors are not compensating:  $h/e \approx 0.7$ , for the ATLAS central barrel calorimeter, is typical.

A more versatile approach to compensation is provided by a *dual-readout calorimeter*, in which the signal is sensed by two readout systems with highly contrasting  $h/e$ . Although the concept is more than two decades old [140], it has only recently been implemented by the DREAM collaboration [141]. The test beam calorimeter consisted of copper tubes, each filled with scintillator and quartz fibers. If the two signals  $Q$  and  $S$  (quartz and scintillator) are both normalized to electron response, then for each event Eq. (1.39) takes the form:

$$\begin{aligned} Q &= E [f_{em} + h/e|_Q (1 - f_{em})] \\ S &= E [f_{em} + h/e|_S (1 - f_{em})] \end{aligned} \quad (1.40)$$

These equations are linear in  $1/E$  and  $f_{em}$ , and are easily solved to obtain estimators of the *corrected* energy and  $f_{em}$  for each event. Both are subject to resolution effects, but effects due to fluctuations in  $f_{em}$  are eliminated. The solution for the corrected energy is given by [135]:

$$E = \frac{RS - Q}{R - 1}, \text{ where } R = \frac{1 - h/e|_Q}{1 - h/e|_S} \quad (1.41)$$

$R$  is the energy-independent slope of the event locus on a plot of  $Q$  vs  $S$ . It can be found either from the fitted slope or by measuring  $\pi/e$  as a function of  $E$ .

Although the usually-dominant contribution of the  $f_{em}$  distribution to the resolution can be minimized by compensation or the use of dual calorimetry, there remain significant contributions to the resolution:

1. Incomplete corrections for leakage, differences in light collection efficiency, and electronics calibration.

## 52 1. Detectors at accelerators

2. Readout transducer shot noise (usually photoelectron statistics), plus electronic noise.
3. Sampling fluctuations. Only a small part of the energy deposit takes place in the scintillator or other sensor, and that fraction is subject to large fluctuations. This can be as high as  $40\%/\sqrt{E}$  (lead/scintillator). It is even greater in the Fe/scint case because of the very small sampling fraction (if the calorimeter is to be compensating), and substantially lower in a U/scint calorimeter. It is obviously zero for a homogeneous calorimeter.
4. Intrinsic fluctuations. The many ways ionization can be produced in a hadronic shower have different detection efficiencies and are subject to stochastic fluctuations. In particular, a very large fraction of the hadronic energy ( $\sim 20\%$  for Fe/scint,  $\sim 40\%$  for U/scint) is “invisible,” going into nuclear dissociation, thermalized neutrons, etc. The lost fraction depends on readout—it will be greater for a Cherenkov readout, less for a heterogeneous active medium such as organic scintillator.

Except in a sampling calorimeter especially designed for the purpose, sampling and intrinsic resolution contributions cannot be separated. This may have been best studied by Drews *et al.* [142], who used a calorimeter in which even- and odd-numbered scintillators were separately read out. Sums and differences of the variances were used to separate sampling and intrinsic contributions.

The fractional resolution can be represented by

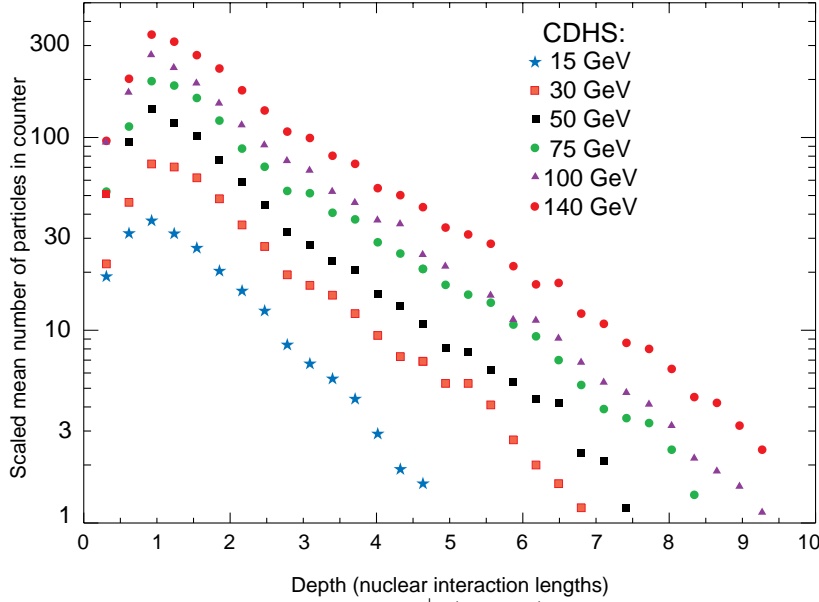
$$\frac{\sigma}{E} = \frac{a_1(E)}{\sqrt{E}} \oplus \left| 1 - \frac{h}{e} \right| \left( \frac{E}{E_1} \right)^{1-\ell} \quad (1.42)$$

The coefficient  $a_1$  is expected to have mild energy dependence for a number of reasons. For example, the sampling variance is  $(\pi/e)E$  rather than  $E$ .  $(E/E_1)^{1-\ell}$  is the parameterization of  $\sigma_{fem}$  discussed above. At a time when data were of lower quality, a plot of  $(\sigma/E)^2$  vs  $1/E$  was apparently well-described by a straight line (constant  $a_1$ ) with a finite intercept—the square of the right term in Eq. (1.42), then called “the constant term.” Modern data show the slight downturn [131].

After the first interaction of the incident hadron, the average longitudinal distribution rises to a smooth peak. The peak position increases slowly with energy. The distribution becomes a reasonably exponential after several interaction length. Examples from the CDHS magnetized iron-scintillator sandwich calorimeter test beam calibration runs [143] are shown in Fig. 1.23. Proton-induced cascades are somewhat shorter and broader than pion-induced cascades [136]. A gamma distribution fairly well describes the longitudinal development of an EM shower, as discussed in Sec. 30.5. Following this logic, Bock *et al.* suggested that the profile of a hadronic cascade could be fitted by the sum of two gamma distributions, one with a characteristic length  $X_0$  and the other with length  $\lambda_I$  [144]. Fits to this 4-parameter function are commonly used, *e.g.*, by the ATLAS Tilecal collaboration [136]. If the interaction point is not known (the usual case), the distribution must be convoluted with an exponential in the interaction length of the incident particle. Adragna *et al.* give an analytic form for the convoluted function [136].

The transverse energy deposit is characterized by a central core dominated by EM cascades, together with a wide “skirt” produced by wide-angle hadronic interactions [145].

The CALICE collaboration has tested a “tracking” calorimeter (AHCAL) with highly granular scintillator readout [132]. Since the



**Figure 1.23:** Mean profiles of  $\pi^+$  (mostly) induced cascades in the CDHS neutrino detector [143].

position of the first interaction is observed, the average longitudinal and radial shower distributions are obtained.

### 1.10. Superconducting magnets for collider detectors

Revised September 2011 by A. Yamamoto (KEK); revised October 2001 by R.D. Kephart (FNAL)

**1.10.1. Solenoid Magnets:** In all cases SI unit are assumed, so that the magnetic field,  $B$ , is in Tesla, the stored energy,  $E$ , is in joules, the dimensions are in meters, and  $\mu_0 = 4\pi \times 10^{-7}$ .

The magnetic field ( $B$ ) in an ideal solenoid with a flux return iron yoke, in which the magnetic field is  $< 2$  T, is given by

$$B = \mu_0 n I \quad (1.43)$$

where  $n$  is the number of turns/meter and  $I$  is the current. In an air-core solenoid, the central field is given by

$$B(0,0) = \mu_0 n I \frac{L}{\sqrt{L^2 + 4R^2}}, \quad (1.44)$$

where  $L$  is the coil length and  $R$  is the coil radius.

In most cases, momentum analysis is made by measuring the circular trajectory of the passing particles according to  $p = mv\gamma = qrB$ , where  $p$  is the momentum,  $m$  the mass,  $q$  the charge,  $r$  the bending radius. The sagitta,  $s$ , of the trajectory is given by

$$s = qB\ell^2/8p, \quad (1.45)$$

where  $\ell$  is the path length in the magnetic field. In a practical momentum measurement in colliding beam detectors, it is more effective to increase the magnetic volume than the field strength, since

$$dp/p \propto p/B\ell^2, \quad (1.46)$$

where  $\ell$  corresponds to the solenoid coil radius  $R$ . The energy stored in the magnetic field of any magnet is calculated by integrating  $B^2$  over all space:

$$E = \frac{1}{2\mu_0} \int B^2 dV \quad (1.47)$$

## 54 1. Detectors at accelerators

If the coil thin, (which is the case if it is to superconducting coil), then

$$E \approx \left( B^2/2\mu_0 \right) \pi R^2 L . \quad (1.48)$$

For a detector in which the calorimetry is outside the aperture of the solenoid, the coil must be thin in terms of radiation and absorption lengths. This usually means that the coil is superconducting and that the vacuum vessel encasing it is of minimum real thickness and fabricated of a material with long radiation length. There are two major contributors to the thickness of a thin solenoid:

**Table 1.9:** Progress of superconducting magnets for particle physics detectors.

Experiment	Laboratory	$B$ [T]	Radius [m]	Length [m]	Energy [MJ]	$X/X_0$	$E/M$ [kJ/kg]
TOPAZ*	KEK	1.2	1.45	5.4	20	0.70	4.3
CDF*	Tsukuba/Fermi	1.5	1.5	5.07	30	0.84	5.4
VENUS*	KEK	0.75	1.75	5.64	12	0.52	2.8
AMY*	KEK	3	1.29	3	40	†	
CLEO-II*	Cornell	1.5	1.55	3.8	25	2.5	3.7
ALEPH*	Saclay/CERN	1.5	2.75	7.0	130	2.0	5.5
DELPHI*	RAL/CERN	1.2	2.8	7.4	109	1.7	4.2
ZEUS*	INFN/DESY	1.8	1.5	2.85	11	0.9	5.5
H1*	RAL/DESY	1.2	2.8	5.75	120	1.8	4.8
BaBar*	INFN/SLAC	1.5	1.5	3.46	27	†	3.6
D0*	Fermi	2.0	0.6	2.73	5.6	0.9	3.7
BELLE*	KEK	1.5	1.8	4	42	†	5.3
BES-III	IHEP	1.0	1.475	3.5	9.5	†	2.6
ATLAS-CS	ATLAS/CERN	2.0	1.25	5.3	38	0.66	7.0
ATLAS-BT	ATLAS/CERN	1	4.7–9.75	26	1080	(Toroid)†	
ATLAS-ET	ATLAS/CERN	1	0.825–5.35	5	$2 \times 250$	(Toroid)†	
CMS	CMS/CERN	4	6	12.5	2600	†	12

\* No longer in service

† EM calorimeter is inside solenoid, so small  $X/X_0$  is not a goal

- 1) The conductor consisting of the current-carrying superconducting material (usually NbTi/Cu) and the quench protecting stabilizer (usually aluminum) are wound on the inside of a structural support cylinder (usually aluminum also). The coil thickness scales as  $B^2 R$ , so the thickness in radiation lengths ( $X_0$ ) is

$$t_{\text{coil}}/X_0 = (R/\sigma_h X_0) \left( B^2/2\mu_0 \right) , \quad (1.49)$$

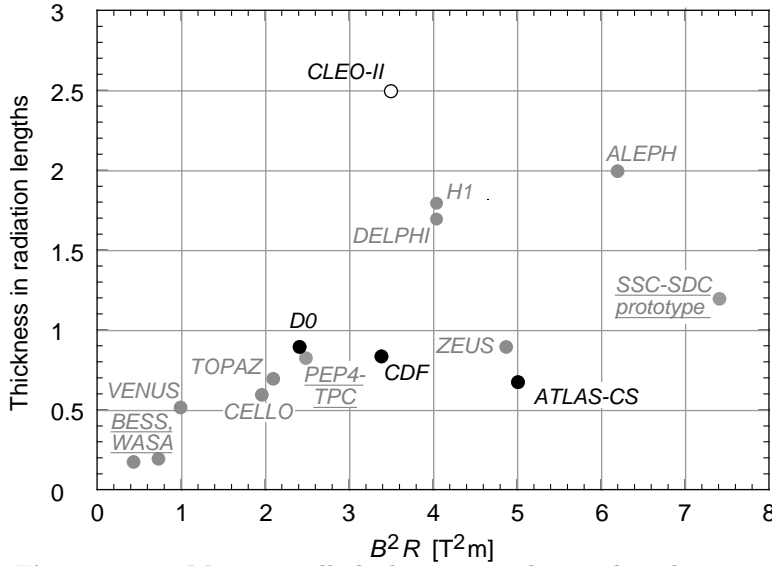
where  $t_{\text{coil}}$  is the physical thickness of the coil,  $X_0$  the average radiation length of the coil/stabilizer material, and  $\sigma_h$  is the hoop stress in the coil [155].  $B^2/2\mu_0$  is the magnetic pressure. In large detector solenoids, the aluminum stabilizer and support cylinders dominate the thickness; the superconductor (NbTi/Cu) contributes a smaller fraction. The main coil and support cylinder components typically contribute about 2/3 of the total thickness in radiation lengths.

- 2) Another contribution to the material comes from the outer cylindrical shell of the vacuum vessel. Since this shell is susceptible to buckling collapse, its thickness is determined by the diameter, length and the modulus of the material of which it is fabricated.

The outer vacuum shell represents about 1/3 of the total thickness in radiation length.

### 1.10.2. Properties of collider detector magnets:

The physical dimensions, central field stored energy and thickness in radiation lengths normal to the beam line of the superconducting solenoids associated with the major collider are given in Table 1.9 [154]. Fig. 1.24 shows thickness in radiation lengths as a function of  $B^2R$  in various collider detector solenoids.



**Figure 1.24:** Magnet wall thickness in radiation length as a function of  $B^2R$  for various detector solenoids. Gray entries are for magnets no longer in use, and entries underlined are not listed in Table 1.9. Open circles are for magnets not designed to be “thin.” The SSC-SDC prototype provided important R&D for LHC magnets.

The ratio of stored energy to cold mass ( $E/M$ ) is a useful performance measure. It can also be expressed as the ratio of the stress,  $\sigma_h$ , to twice the equivalent density,  $\rho$ , in the coil [155]:

$$\frac{E}{M} = \frac{\int (B^2/2\mu_0) dV}{\rho V_{\text{coil}}} \approx \frac{\sigma_h}{2\rho} \quad (1.50)$$

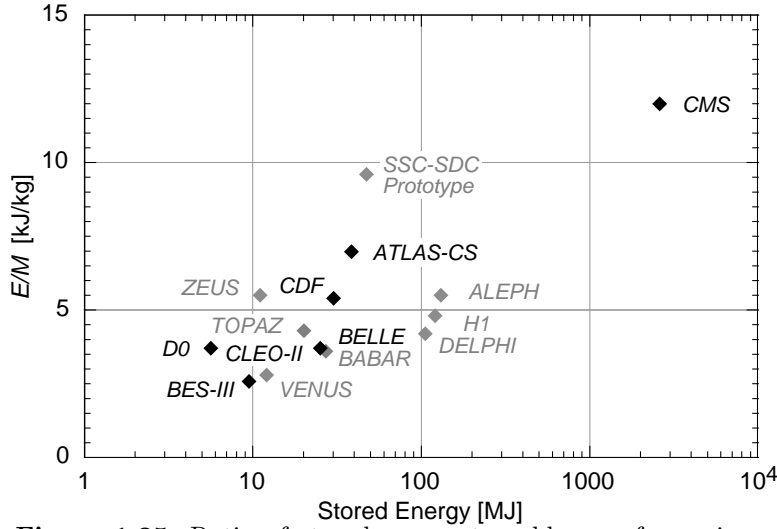
The  $E/M$  ratio in the coil is approximately equivalent to  $H$ ,\* the enthalpy of the coil, and it determines the average coil temperature rise after energy absorption in a quench:

$$E/M = H(T_2) - H(T_1) \approx H(T_2) \quad (1.51)$$

where  $T_2$  is the average coil temperature after the full energy absorption in a quench, and  $T_1$  is the initial temperature.  $E/M$  ratios of 5, 10, and 20 kJ/kg correspond to  $\sim 65$ ,  $\sim 80$ , and  $\sim 100$  K, respectively. The  $E/M$  ratios of various detector magnets are shown in Fig. 1.25 as a function of total stored energy. One would like the

\* The enthalpy, or heat content, is called  $H$  in the thermodynamics literature. It is not to be confused with the magnetic field intensity  $B/\mu$ .

cold mass to be as small as possible to minimize the thickness, but temperature rise during a quench must also be minimized. An  $E/M$  ratio as large as 12 kJ/kg is designed into the CMS solenoid, with the possibility that about half of the stored energy can go to an external dump resistor. Thus the coil temperature can be kept below 80 K if the energy extraction system work well. The limit is set by the maximum temperature that the coil design can tolerate during a quench. This maximum local temperature should be  $<130$  K (50 K + 80 K), so that thermal expansion effects in the coil are manageable.



**Figure 1.25:** Ratio of stored energy to cold mass for major detector solenoids. Gray indicates magnets no longer in operation.

### 1.10.3. Toroidal magnets:

Toroidal coils uniquely provide a closed magnetic field without the necessity of an iron flux-return yoke. Because no field exists at the collision point and along the beam line, there is, in principle, no effect on the beam. On the other hand, the field profile generally has  $1/r$  dependence. The particle momentum may be determined by measurements of the deflection angle combined with the sagitta. The deflection (bending) power  $BL$  is

$$BL \approx \int_{R_i}^{R_0} \frac{B_i R_i dR}{R \sin \theta} = \frac{B_i R_i}{\sin \theta} \ln(R_0/R_i) , \quad (1.52)$$

where  $R_i$  is the inner coil radius,  $R_0$  is the outer coil radius, and  $\theta$  is the angle between the particle trajectory and the beam line axis. The momentum resolution given by the deflection may be expressed as

$$\frac{\Delta p}{p} \propto \frac{p}{BL} \approx \frac{p \sin \theta}{B_i R_i \ln(R_0/R_i)} . \quad (1.53)$$

The momentum resolution is better in the forward/backward (smaller  $\theta$ ) direction. The geometry has been found to be optimal when  $R_0/R_i \approx 3-4$ . In practical designs, the coil is divided into 6–12 lumped coils in order to have reasonable acceptance and accessibility. This causes the coil design to be much more complex. The mechanical structure needs to sustain the decentering force between adjacent coils, and the peak field in the coil is 3–5 times higher than the useful magnetic field for the momentum analysis [153].



### 1.11. Measurement of particle momenta in a uniform magnetic field [156,157]

The trajectory of a particle with momentum  $p$  (in GeV/c) and charge  $ze$  in a constant magnetic field  $\vec{B}$  is a helix, with radius of curvature  $R$  and pitch angle  $\lambda$ . The radius of curvature and momentum component perpendicular to  $\vec{B}$  are related by

$$p \cos \lambda = 0.3 z B R , \quad (1.54)$$

where  $B$  is in tesla and  $R$  is in meters.

The distribution of measurements of the curvature  $k \equiv 1/R$  is approximately Gaussian. The curvature error for a large number of uniformly spaced measurements on the trajectory of a charged particle in a uniform magnetic field can be approximated by

$$(\delta k)^2 = (\delta k_{\text{res}})^2 + (\delta k_{\text{ms}})^2 , \quad (1.55)$$

where  $\delta k$  = curvature error

$\delta k_{\text{res}}$  = curvature error due to finite measurement resolution

$\delta k_{\text{ms}}$  = curvature error due to multiple scattering.

If many ( $\geq 10$ ) uniformly spaced position measurements are made along a trajectory in a uniform medium,

$$\delta k_{\text{res}} = \frac{\epsilon}{L'^2} \sqrt{\frac{720}{N+4}} , \quad (1.56)$$

where  $N$  = number of points measured along track

$L'$  = the projected length of the track onto the bending plane

$\epsilon$  = measurement error for each point, perpendicular to the trajectory.

If a vertex constraint is applied at the origin of the track, the coefficient under the radical becomes 320.

The contribution due to multiple Coulomb scattering is approximately

$$\delta k_{\text{ms}} \approx \frac{(0.016) (\text{GeV}/c) z}{L p \beta \cos^2 \lambda} \sqrt{\frac{L}{X_0}} , \quad (1.58)$$

where  $p$  = momentum (GeV/c)

$z$  = charge of incident particle in units of  $e$

$L$  = the total track length

$X_0$  = radiation length of the scattering medium (in units of length; the  $X_0$  defined elsewhere must be multiplied by density)

$\beta$  = the kinematic variable  $v/c$ .

More accurate approximations for multiple scattering may be found in the section on Passage of Particles Through Matter (Sec. 30 of this *Review*). The contribution to the curvature error is given approximately by  $\delta k_{\text{ms}} \approx 8 s_{\text{plane}}^{\text{rms}} / L^2$ , where  $s_{\text{plane}}^{\text{rms}}$  is defined there.